ERIC Document Resume | See Instruction Sheet | DATE OF RESUME | DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER ED001897

2. ORIGINATOR CONTROL NO(S). N.A.

3. IDENTIFICATION NO. CR - 1166

DO NOT WRITE IN SPACE BELOW

4. AUTHOR(S): Stott, Leland H. and Ball, Rachel S.

MICROFILM COPY AVAILABLE YES ☒ NO ☐

5. TITLE OF DOCUMENT ABSTRACTED: Evaluation of Infant and Preschool Mental Tests

COPYRIGHT GRANTED YES ☐ NO ☒

COPYRIGHT RELEASE FILED YES ☐ NO ☐

6. PUBLICATION TITLE: N.A.

TYPE OF REPORT
- PROPOSAL ☐
- IN-PROGRESS ☐
- FINAL ☒
- FOLLOW-UP ☐

7. EDITOR(S): N.A.

8. PLACE OF ORIGIN
ORG. The Merrill-Palmer Institute
ST. CITY, STATE Detroit, Michigan
TELE. EXT.
INST/LOC CODE

9. PUBLISHER OR SOURCE Cooperative Research Program, Office of Education, Dept. of Health, Education, & Welfare, Wash., D.C.

OTHER ERIC REFERENCES:

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	1963	281

ERIC SATELLITE CODE

B. INDEXING DATA

14. SUBJECT AREA

15. TYPE OF ENDEAVOR Report

16. INDEXING TERMS (6-12)	intelligence theory	factor analysis	preschool
preschool tests	mental development	test content	mental tests
mental ability	determinants of intelligence	tests and measurements	evaluation
test evaluation	mental tests	infants	

17. ABSTRACT (MAX. 250 WORDS)

An investigation and evaluation of total infant and preschool mental testing situations and some conclusions as to the current needs in this area are presented.

These approaches are made to the problem. First, the theoretical thinking on the nature of intelligence and how it develops in children is traced. There has been disagreement over the nature of intelligence, whether it consists of many specific abilities or one general ability. With the development of the factor-analytic methods, however, general intelligence was found to break down into many different but related abilities. After a period of controversy over the determinants of intelligence, it is now accepted that intelligence develops through the interaction of organism and environment, and that intelligence develops in a qualitative, as well as a quantitative sense.

The second approach was a questionnaire survey of the uses being made of mental tests for young children. Information is given on the frequency with which each of the various tests is used and for what purposes. Special values as well as inadequacies were reported for each of the widely used tests, and many suggestions for improvement were offered.

The third approach was a statistical analysis of an accumulated mass of test results. The findings indicate that, even as young as three months, children are capable of "thinking" processes. There is great variety in factor content among different tests and for different age levels. Although other inadequacies are described, there is no evidence that the results of the infant tests show less valid discrimination of ability than do tests for older age levels.

ERIC No.

18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION

20. DIVISION OR DEPT.

21. CONTRACT NO.

PROJECT NO.

22. TOTAL COST OF PROJECT

23. DURATION OR PROJECT

U.S. Department of Health, Education, and Welfare
Office of Education

5

ERIC Form 1

ED001897

EVALUATION OF INFANT AND PRESCHOOL MENTAL TESTS

REPRINTED BY PERMISSION
U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN PRINTED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

EVALUATION OF INFANT AND PRESCHOOL MENTAL TESTS

Cooperative Research Project No. 1166

Leland H. Stott
The Merrill-Palmer Institute

Rachel Stutsman Ball
Arizona State University

The Merrill-Palmer Institute
Detroit, Michigan
1963

The research reported herein was supported through the Cooperative Research Program of the Office of Education, United States Department of Health, Education and Welfare.

Acknowledgements

Although we, the authors, assume full responsibility for the form and the content of this report, we do wish gratefully to acknowledge the help and participation of many individuals in the various aspects of the project. First of all we acknowledge with gratitude the financial support of the Cooperative Research Branch of the Office of Education, United States Department of Health, Education, and Welfare, and especially do we appreciate the help, the counsel and the support throughout the thirty months duration of the project of Dr. Robert Beezer of that office. And likewise without the encouragement and support of President Pauline P. W. Knapp and Vice President William W. McKee of the Merrill-Palmer Institute, the contracting institution, the project could not have been carried out.

Special recognition and our sincere thanks are extended to our consultants, Professor Raymond Cattell of the University of Illinois and Professors J. P. Guilford and Phillip Merrifield of the University of Southern California. Although we, the two investigators, assume full responsibility for any shortcomings this report may reveal, without the guidance and the help of our consultants, many of the unanticipated technical problems and difficulties which arose in the course of the research could not have been satisfactorily resolved.

We also acknowledge with gratitude the technical advice, and the readily available services of our programmer and computer expert, Dr. James Ball of the University of California at San Diego. His willingness to come to our assistance and carry out additional and repeated analyses at a moments notice was greatly appreciated.

To Drs. Nancy Bayley who contributed test records from her files, and offered her suggestions of factor names in the analysis of the California Scale, and to Drs. Marie Skodak and Libby Bass who also contributed test protocols for our analyses, we extend our thanks.

It is obvious that space will not allow us to mention individually the many users of tests (psychologists and clinicians, heads of departments and directors of agencies) who answered our questionnaire and made so many additional comments and suggestions by note or separate letter. To each of our respondents we are very grateful.

A number of others contributed importantly to the progress of the project along the way, and to its final completion. To Leon Wells, Arthur Beirlockway, Fred Stevens, Howard Shapiro, Dennis Campbell, Nancy Eddy, and Baqar Husaini who performed much of the drudgery of the project, especially do we offer our sincere thanks.

Finally we want especially to recognize the invaluable services of our secretary, Mrs. Irene Zak. She assisted in many ways throughout the duration of the project but particularly important was her

EDoo 1897

patient piecing together of preliminary drafts and her painstaking typing of the final copy.

CONTENTS

Page

Acknowledgments iii

List of Tables x

I

INTRODUCTION 1

II

INTELLIGENCE,
ITS NATURE AND DEVELOPMENT 7

A Word with Many Meanings 7

Points of View about the Nature of Intelligence and Its Determining Factors 10
- The Assumption of Inherently Fixed Intelligence 10
- Predeterminism 13
- Environment as a Factor in Development 16
- The Controversy 18
- The Interaction Point of View 31

Points of View Regarding the "Structure" of Intelligence 36
- The Concept of General Intelligence 36
- Concepts of Multiple Mental Abilities 38
- A Three-Dimensional Model of Intellect 40

The Problem of Developmental Change in Mentality 44
- The Assumption of Quantitative Change 46
- Binet on Mental Developmental Change 47
- Genetically Endowed Potentiality 48
- The Theory of Quantitative Change and the Lack of Predictive Value in Infant Tests 50

The Problem of Developmental Change in Mentality (Continued) Page

Development as Qualitative Change 54
Gesell's View of Developmental Change 56
Piaget's Epigenetic Theory 59
Other Conceptions of Intelligence which Imply Qualitative Change 69
Developmental Sequences and the Measurement Problem 75

Summary - Our Present Theoretical Orientation 77

III

INTELLIGENCE TESTING AT THE INFANT AND PRESCHOOL LEVELS 84

The Development and Evaluation of Tests 84
Problems of Test Construction 84
The Binet Tests 85
The Gesell Developmental Schedules 91
The Merrill-Palmer Scale of Mental Tests 94
The Minnesota Preschool Scale 100
The Goodenough Draw-a-Man Test 101
The California First Year Mental Scale 102
The Cattell Infant Intelligence Scale 103
The Northwestern Infant Intelligence Test 105
Early Performance Scales 106
The Leiter International Performance Scale 107
The Full Range Picture Vocabulary Test 108
The Griffiths Mental Development Scale 109
The A. Brenner Gestalt Test 110
A New Revision of the California First Year Scale 110

Published Studies which Involved the Testing of Infants and Preschool Children 112
Studies Concerned with the Nature of Early Intelligence and the Factors in Its Development 112
Use of Intelligence Tests in Early Childhood for Research Concerning Other Psychological Variables 118

IV

QUESTIONNAIRE SURVEY OF THE CURRENT SITUATION WITH RESPECT TO INFANT AND PRESCHOOL MENTAL TESTING — Page 123

Results of the Survey 123
    Services to Young Children and Their Parents 124
    Types of Cases Handled by the Respondents 125
    Infant and Preschool-Age Scales in Use 126

Evaluative Statements 126
    The Stanford-Binet 126
    The Goodenough Draw-a-Man 127
    The Wechsler Intelligence Test for Children 128
    The Gesell Schedule 128
    The Cattell Infant Scale 129
    The Ammons Full Range Picture Vocabulary Test 129
    The Merrill-Palmer Scale 130

Limitations and Suggestions for Improvement 130

V

FACTOR ANALYSIS OF INFANT AND PRESCHOOL TEST RESULTS 133

Problems Encountered 138
    The Problem of Experimenta Dependencies 138
    Determination of the Number of Factors to Be Extracted 142
    Factor Identification and Interpretation 143

General Procedure 144
    Statistical Procedure 144
    Interpretation Procedures 146
    The Choice of a Rotation for Factor Interpretation 148

The Cattell Intelligence Test 149
    The Three Months Level 149
    Cattell Scale, Six Months Level 152

California First Year Scale 156
    Six Months Level 157
    Twelve Months Level 161

Page

Gesell Developmental Schedule 165
    Gesell, Six Months Level 166
    Gesell, Twelve Months Level 170

The Merrill-Palmer Scale 175
    Twenty-four Months Age Level 176
    Thirty Months Level 181
    The Thirty-six Months Level 184
    Merrill-Palmer, Forty-eight Months Level 190
    The Fifty-four Months Level 195

The Stanford-Binet (Form L) Scale 198
    The Three Year Age Level 199
    The Forty-eight Months Level 204
    The Sixty Months Level 207

Discussion of Factor Findings 210
    Comparison of Factor Meaning 217

VI

SUMMARY AND CONCLUSIONS 223

    Theories and Points of View 223
    Uses Currently Being Made of Infant and Preschool Mental Tests 225
    Factor Analysis Infant and Preschool Mental Test Content 227
    General Conclusions 229

REFERENCES 232

APPENDIX 245
    Tables 247
    Survey of Infant and Preschool Mental Testing (Questionnaire) 270

List of Tables

1. Types of Services Offered Children and Their Families, with Percentage Frequencies
2. Types of Cases Most Frequently Reported by 217 Respondents
3. The Fifteen Most Frequently Used Infant and Preschool Age Mental Test Scales, As Reported by 217 Respondents
4. Mental Tests, with age Levels, Used in the Fourteen Factor Analysis Studies
5. Factor Analysis, Cattell Infant Scale, Age Level, 3 Months
6. Factor Analysis, Cattell Infant Scale, Age Level, 6 Months
7. Factor Analysis, California First Year Mental Scale, Age Level, 6 Months
8. Factor Analysis, California First Year Mental Scale, Age Level, 12 Months
9. Factor Analysis, Gesell Developmental Schedule, Age Level, 6 Months
10. Factor Analysis, Gesell Developmental Schedule, Age Level, 12 Months
11. Factor Analysis, Merrill-Palmer Scale, Age Level, 24 Months
12. Factor Analysis, Merrill-Palmer Scale, Age Level, 30 Months
13. Factor Analysis, Merrill-Palmer Scale, Age Level, 36 Months
14. Factor Analysis, Merrill-Palmer Scale, Age Level, 48 Months
15. Factor Analysis, Merrill-Palmer Scale, Age Level, 54 Months
16. Factor Analysis, Stanford-Binet, Form L, Age Level, 3 Years
17. Factor Analysis, Stanford-Binet, Form L, Age Level, 48 Months
18. Factor Analysis, Stanford-Binet, Form L, Age Level, 60 Months
19. Factor Intercorrelations - Biquartimin Rotations
20. Summary of Factor Interpretations, Allowing Comparisons of Tests and of Levels as to Interpreted Factor Content

I

INTRODUCTION

In very recent years developmental psychologists have been showing an increased interest in the earliest phases of human development. The period of infancy has always been an area of interest and concern, but in the early days of the study of Child Development the main focus was upon the baby's physical well being - his care, his nutrition, his rhythms of sleep and wakefulness, his patterns of feeding and elimination. Among the vital concerns of pediatricians and the child care specialists of a quarter of a century ago were such issues as breast versus bottle feeding, self-regulation versus clock scheduling of the infant's feeding periods, and the problems of "toilet training." Today the focus of interest seems to have shifted in the direction of what the baby is like psychologically at birth - his temperamental nature, his "intellective" functions and when they appear, and the extremely rapid changes that take place in them during the early weeks and months of infancy. There is a lively and growing interest in cognitive development, in the role of stimulation in early intellective development and in the possibilities of facilitating optimal development, without pressure, but through an awareness of certain "critical periods" of development, being able to provide appropriate conditions of stimulation.

Our conceptions of the nature of mentality and of the processes of intellectual development are undergoing modification in terms of these newer interests. Whereas formerly, intelligence was generally conceived of as a general, unitary, pervasive quality which, in terms of amount, characterized the individual throughout life, it is now more frequently conceived as interrelated abilities or functions having to do with one's coming to know and to cope with one's environment. The idea that intelligence is "fixed" by heredity, not only in terms of the amount the individual possesses, but also in its rate and pattern of development, is generally being replaced by the view that the environment is also an important determinant in interaction with the organism. Implied also in these newer viewpoints is the idea that developmental changes in mentality are qualitative, as well as quantitative in nature.

Another conceptual change concerning the nature of intelligence, which is coming, and which we feel is overdue, is in relation to the dichotomy of "motor" behavior versus mental or central processes. The infant begins to "know," and to cope with his environment through his motor responses to stimulation. His eyes follow a patch of color. He responds to the sound by another's approaching steps. He kicks and thrashes about with his arms in a random, disorganized manner thus bringing his hands in contact with objects within his reach. These he grasps, mouthes them, shakes and bangs them, and thus "experiences" them with his various senses. It is quite clear, however, that in the beginning these simple, sensorimotor experiences do not involve the awareness of the objects from which the stimulation comes. There could be no central representation of them, as such, at first. But developmental change is rapidly taking place. Soon, with further maturation and experience the child learns to differentiate objects, and to respond to them as objects as he sees them, grasps them, tastes them, hears them fall to the floor. His motor behavior now involves the cognitive awareness of things and people. His functioning with respect to them has become meaningful. All his motor manipulations - his eye-hand, fine motor, and gross motor coordinations appear now to be centrally mediated and are directed toward some kind of adjustment to, or coping with the environment. In that sense, therefore, they are "intellectual" in nature. In short, the motor behavior of the young child is mental behavior.

The presently available and widely used tests and scales for measuring mentality at the infant, and preschool levels, with few exceptions, were devised and standardized in terms of the older conceptions of the nature and development of intelligence. They were generally standardized at these early age levels, upon too few, and, in many instances, poorly selected cases. Nevertheless these tests have proved to be very useful instruments when used with discretion. In general, the present-day users of these scales, although they find them indispensable in their work, are well aware of their defects and inadequacies. They would like to have the scales amplified, modernized and restandardized.

One of these early preschool scales, the Merrill-Palmer Scale of Mental Tests, was developed by one of the present authors, Rachel Stutsman (Ball), some thirty-five years ago. Although it is still in use it is clearly in need of modernization, revision or replacement. It is our ultimate objective to accomplish this task. It was clear from the beginning that much preliminary work was necessary. A first step that appeared to be important was to make a comprehensive evaluation of the present situation with respect to infant and preschool mental testing in this country. The project with which this report is concerned was designed to accomplish such an evaluation.

An evaluation of this nature, as we envisaged it, was to be made from three main approaches. First, the technical and professional literature dealing with the concept of intelligence and its measurement, particularly at the infant and early childhood levels, was reviewed, organized and summarized. This summary review is presented in Parts II and III of this report. Part II deals briefly with some of the main trends of thinking about the nature of intelligence that have predominated in the past and that are now current. Special attention is given to the problem of the nature of developmental change in mentality during infancy and childhood. Part III deals with the development of tests and scales for the measurement of mental ability during these early stages, and with the various uses that have been made of these testing devices.

As a second approach, a questionnaire survey was made of the current situation with respect to the testing of infant, and preschool-age intelligence. The questionnaire was designed to provide an estimate of the relative frequencies with which the various tests are presently being used, and for what specific purposes. The respondents were asked to express their judgments as to the usefulness and adequacy of the tests they are using in relation to their particular needs, to point out any special usefulness they have discovered in the tests, as well as their inadequacies, for particular purposes, and to suggest needed changes in the current testing situation. The results of this survey are presented in Part IV.

Our third approach was to analyze statistically the actual content of five of the most widely used infant and preschool scales, in terms of children's actual performances on each specific test item. The purpose was to determine (1) to what extent these test scales are similar in item-meaning content - whether they involve the same or different ability factors -, and (2) to determine the consistency, or lack of consistency, with which each scale maintains the same meaning content throughout its various age levels.

It was decided that the best available approach to making such comparisons was by means of factor analysis. Since these scales were designed to cover age ranges varying from 12 to 36 months, it was found necessary to solve some 15 separate factor analysis problems. With the help of modern computers and the many new developments in terms of techniques and "programs," this task, which a few years ago would have been a formidable one, was readily accomplished. This third phase of our project is the subject of Part V of this report.

Although the two investigators worked in close cooperation on all phases of the study, Leland H. Stott assumed the main task of reviewing the literature and conducting the questionnaire survey, while Rachel Stutsman Ball bore the main burden of the factor analyses. We share equally in the responsibility for this report.

II

INTELLIGENCE, ITS NATURE AND DEVELOPMENT

A considerable portion of the psychological literature during the first half of the present century has been concerned with the problem of "intelligence," its general nature and its measurement. It has long been considered a matter of prime importance to the understanding of human behavior. Children obviously differ tremendously in their readiness and ability to learn and in their rate of mental developmental progress. The nature of the factor or factors underlying these differences has become a matter of theoretical interest, and the assessment and the equitable social management of individual differences in mental ability continue to be matters of great practical importance. Before beginning a historical account of the predominant ways in which intellectual ability and its development have been conceived, a brief review of some of the common, divergent meanings with which the term "intelligence" is used by present-day psychologists is presented.

A Word With Many Meanings

"Intelligence" is one of the many common-parlence terms taken over by Psychology and given "scientific" or technical meaning. This fact alone makes for difficulty in arriving at any degree of consistency or preciseness in its usage. What is even more difficult is the fact that in Psychology the term is used with various meaning. "Intelligence" is often defined as ability, i.e., what the subject can do, his level of performance at the time (Stoddard, 1943, p. 4). In this sense intelligence changes; it increases with age as the child develops. The most common symbol indicating the attained level of ability at any point during this developmental period, of course, is the MA (mental age).

Another equally legitimate "technical" use of the term intelligence is to let it refer, not to ability or abilities, *per se*, but to the inherited potentiality for the development of those abilities. In this sense "intelligence is fixed" by the genes at conception.

Still another very common way of conceiving of intelligence is in terms of the child's rate of mental development as measured by the ratio of mental age to chronological age (IQ). The higher the ratio, the "brighter" or more intelligent the child is said to be. Here again, the focus is not on what the child *can do* -- on his abilities as such. Neither is it a measure of potentiality, except in a very indirect way, i.e., on the very tenuous assumption that the greater the potentiality, the higher the rate of progress toward the achievement of that potentiality. The focus is upon the rate at which the child's mentality is developing.

There are a number of less frequent meanings attached to the term intelligence. One is that it is a *function* of a living organism "in much the same sense that maneuverability and speed are functions of an airplane" (Munn, 1946, p. 410). In the view of the present writer, it would be more accurate to say that intelligence is a generalized attribute of individual functioning in general, rather than a function, just as maneuverability and speed are attributes of the functioning of an airplane. It would seem that this is what Munn really means as he suggests in his final definition of intelligence. "Intelligence is flexibility and versatility in the use of symbolic processes" (1946, p. 411). The "use of symbolic processes" is a category of functioning, and "flexibility" and "versatility" are descriptive of the overall quality or attribute called "intelligence." In this sense, intelligence means effectiveness, facility, flexibility, versatility, general adequacy of functioning in life situations.

Other usages of the term make less sense. For example, to say that intelligence is "native ability" in a strict sense, limits its reference only to what a newborn baby can do. Native means to be born with. As an example of this type of unprecise definition, Burt, et. al. wrote: "By intelligence, the psychologist understands inborn, all-around intellectual ability.... It is inherited, or at least innate" (1934). Very little of intellectual ability is innate. The child is not born with that ability, only with the potentiality for developing that ability. Undoubtedly these authors meant to refer to the potentiality, rather than ability per se, as being "inborn," "innate."

More recent usages define intelligence, not as ability to do, not as a genetic capacity to develop ability, not as an attribute or quality of functioning, but rather in terms of actual processes of central functioning. "Intelligence is conceived of as central operations for the processing of information" (Hunt, 1961, p. 337). These central processes, moreover, are conceived as "hierarchically arranged within the intrinsic portions of the cerebrum" and as "approximately analogous to the strategies for information processing and action with which electronic computers are programmed" (p. 362).

Points of View About the Nature of Intelligence and Its Determining Factors

Some of the divergence in meaning with which the term intelligence is used comes naturally from different broad conceptualizations and points of view which have prevailed from time to time in the past concerning the nature of human abilities and how they develop.

The Assumption of Inherently Fixed Intelligence. The general idea that intelligence is genetically "fixed" as a pervading characteristic of the individual has had perhaps most general acceptance in the past by both theoretical writers and makers of tests. The idea undoubtedly had its origin in Darwin's (1869) theories of biological evolution and it gained impetus from the research and writings of Sir Francis Galton on mental inheritance (1869, 1833). As a result of his studies of eminent men in Great Britain, Galton became convinced that intellectual genius is inherited; and he set about with energy to devise tests of simple sensory and motor functions, immediate memory, etc. for assessing this inherited capacity and for demonstrating individual differences. Many thousands of individuals were given these tests in Galton's laboratory, and although wide individual differences were demonstrated, the outcome of his project was disappointing in that his measures showed little relationship with other criteria of intelligence. Galton's ultimate objective, of course, was to set up a program of race improvement in which the intellectually gifted would be identified early and encouraged to reproduce and thus pass on their genius. His simple tests, however, proved not to be an effective means of differentiating the high achievers from the low achievers.

Alfred Binet began his test construction project under different motivation and with quite different objectives than those of Galton. Under his commission from the French Minister of Public Instruction, he began his study of the nature of school retardation and the task of devising a means of assessing various degrees of mental retardation in children. Binet and his co-worker, Simon, published their first scale of mental ability in 1905. Unlike Galton, they looked for tests of the more complex psychological functions such as "judgment, otherwise called good sense, practical sense, initiative, the faculty of adapting oneself to circumstances" (1916, p. 42). And, unlike Galton's tests, their scale proved to be effective in differentiating children who achieved success in school from those who did not.

Binet, however, did not hold to the view that intelligence is genetically fixed. On the contrary, he was vehemently opposed to "the deplorable verdict that the intelligence of an individual is a fixed quantity." (Binet, 1909, p. 54). His tests, nevertheless, furnished evidence which was generally interpreted as support for the assumption of inherited intelligence.

Certain leading figures in American psychology who played a prominent part in the development of the testing movement also lent strong support to the concept of genetically determined intelligence. H. H. Goddard was the first to introduce the Binet-Simon scale, and to make use of it, in this country. In 1906 a psychological laboratory was set up at the Training School for Feebleminded at Vineland, New Jersey with Goddard as its director. He had seen something of the usefulness of the Binet scale in France as it applied to "feeble minded" children, and in 1908 he began using it at Vineland. His first task to was translate the scale into English and to adapt it generally to American conditions. Goddard's work eventuated finally in a restandardization of the scale on 2,000 American children. This revision was published in 1910.

Based on his experiences at Vineland, Goddard had little faith in the educability of the feebleminded. Being strongly oriented in favor of heredity as the basis of individual differences in intelligence, he became a strong advocate of the doctrine of fixed intelligence.

Lewis M. Terman is, of course, closely identified with the mental testing movement in the United States in connection with his Stanford revisions of the Binet scale, the first one in 1916. The only innovation in this revision, however, was the adoption of the "Intelligence Quotient" a concept previously introduced by Stern (1914), as an index of intelligence.

The concept of IQ naturally was a potent reinforcer for the assumption of fixed intelligence. Here was an index of intelligence which showed a wide range of variation in the population, but which, at the same time, remained "relatively constant" from age to age. The IQ, furthermore, showed evidence of validity in that generally children with IQ's below 85 typically fell behind their age group in school achievement while those with high IQ's tended to make better than average progress and to exhibit certain other superior qualities.

Thus psychologists generally, and particularly those concerned with intelligence and its assessment, came rather generally to assume that intelligence as a general capacity is genetically determined and that this capacity remains fixed throughout life.

Predeterminism. Another concept consonant with the assumption of fixed intelligence, but with a different derivation in history, and with implications of quite a different sort, is the concept of predetermination. According to this concept all behavioral development unfolds automatically with the maturation of the anatomical structure of which behavior is a function.

Doctrines of predeterminism, of course, are not of recent origin. The French philosopher, J. J. Rousseau, elaborated his predeterministic theory 200 years ago. He postulated that "all development consists of a series of internally regulated sequential stages which are transformed, one into the other, in conformity with a pre-arranged order and design. According to this conception of development, the only proper role of the environment is avoidance of serious interference with the processes of self-regulation and spontaneous maturation" (Ausubel, 1958, p. 27). As an explanation of the assumed regularity and inevitability with which the predetermined pattern of development in each case is carried out, Rousseau and his followers proposed the idea that the course of individual development parallels that of the species through the various stages of its evolution. A hundred years later this doctrine of "recapitulation" was elaborated and refined by G. Stanley Hall. However, as developmental data accumulated and as theoretical constructs regarding the inter-relationships between the biological and the cultural aspects of development were elaborated, Hall's system of parallelisms was no longer accepted.

A basic tenet of the doctrine of recapitulation, namely that maturation is a fundamental process of anatomical development, is still important in developmental theory. As a biological process in organismic development maturation is well established. However, when maturation is postulated as the sole factor of behavioral, as well as morphological development, we still have predeterminism albeit lacking the elaborative trimmings proposed by G. Stanley Hall.

Arnold Gesell, a student of Hall, is perhaps the strongest contemporary proponent of predetermined development. His great stress on the importance of the "mechanism of maturation" in his model of morphological development has made a strong impact upon current developmental psychology. According to Gesell (1954), the "basic configurations, correlations, and successions of behavior patterns are determined by a process of maturation. The tenacity of this process has been demonstrated by developmental studies of both premature and full-term infants and by experimental studies of monozygotic twins" (In Carmichael, 1954, p. 371).

Clearly, the idea that behavioral development is a product of maturation, and therefore predetermined, encompasses the concept of fixed intelligence. The assumption is that, not only is the intellectual developmental potential of the individual fixed by the genes, but so also is the course, the pattern and the rate of attainment fixed by this potential. It is attained through the "mechanism of maturation" -- a process which is uninfluenced by the individual's encounters with his environment.

It is quite clear also that this assumption was rather generally accepted by psychologists during the era of intelligence test development. The definitions of intelligence that appeared in the Psychology text books during the 1920's and 1930's reflect this double assumption of predetermined development and fixed intelligence. Some typical definitions follow:

> By intelligence, the psychologist understands inborn, all-around, intellectual ability. It is inherited, or at least innate, not due to teaching or training; it is intellectual, not emotional or moral, and remains uninfluenced by industry or zeal (Burt, et al, 1934, pp. 28-29).

> Given equal opportunity and training, two individuals will often be found to differ considerably in the success with which they achieve. There are two factors at work here, namely, native capacity (intelligence) and drive... Intelligence may be broadly defined as the capacity for solving the problems of life (Allport, F. H., 1924, p. 104).

> Intelligence is the capacity to acquire and perfect new modes of adaptation. Since the intelligence of a person is not his stock of habits but his ability to acquire habits, it would seem to follow that this is a matter, not of practice and experience, but of inborn nature (Dashiell, 1928, p. 306).

Environment as a Factor in Mental Development. Most, but not all, of the text book writers of the '20s and '30s apparently adopted the point-of-view implied in the above definitions and so strongly urged by Gesell concerning intellectual development -- that both the upper limit of developmental capacity and its pattern and rate of development are predetermined by heredity. Harvey A. Carr (1926) was one notable exception. In fact he spelled out in some detail a version of the currently widely accepted "interaction" view of mental development. He wrote:

> All experiences of the individual during life are thus organized into a complex but unitary system of reaction tendencies that determine to a large extent the nature of his subsequent activity. The reactive disposition of the individual, i.e., what he does and what he can and cannot do, is a function of his native equipment, of the nature of his previous experiences and of the way in which these have been organized and evaluated (p. 4).
>
> ......Neither is this ability (intelligence) native in the sense that its development is not dependent upon environmental influences... The ability is natively conditioned in that we are born with the capacity to develop in this manner (p. 415).
>
> The assumption that all individual, group, and racial differences of IQ are wholly determined by innate differences of constitution must be discarded because of the fact that the IQ is often altered during the course of development (p. 417).

Perhaps the most widely known controversy to be found in the psychological literature was over this very issue. In 1932, Beth L. Wellman of the Iowa Child Welfare Research Station published the first in a series of studies reporting marked changes in the IQ's of children attributed to changes in environment. This first report is based upon some 3,000 Binet Tests administered to the children of the Iowa longitudinal series. The ages of the children at the time of their first testing ranged between 2 and 14 years. "The mean IQ on the first test for the entire group, 1,333 children, was 110; for the 1,027 cases on the second and third tests, the mean was 119; and for 574 cases on the fourth to seventh tests, 124." (1932a, p. 118). According to the report the greatest gains were made by those children whose initial IQ's were below average; the average gain for them was 28 points on their fifth test. The "average" group had gained 22 points from their first to seventh test, while the superior group made an average gain of 12 points up to their fifth test.

Wellman's explanation for these findings was that,

> "preschool attendance, at least in the laboratories of the Iowa Child Welfare Research Station, causes a rise in IQ, the rise being cumulative from year to year, and sustained throughout the school years when the children are in the environment provided by our University Elementary School. When the same children are home over comparable intervals they fail to gain, although maintaining their higher level" (1932a, p. 124).

Dr. Wellman admitted at the end of her report that "the question still remains whether these increased IQ's are real or inflated."

In her second report however, Wellman (1932b) furnished evidence that the gains in IQ during preschool attendance were not practice effects. She reported that the same children who gained IQ points during the school session failed to gain during a shorter period between tests during the summer vacation. Her results also indicated that during the preschool season greater gains were shown during a 7 to 8 month interval between tests than during a shorter, 5-6 month interval which would be more likely to show practice effects.

In a study of Crissey (1937) the evidence indicated that children who were rated as "normal or dull-normal" on admission to an orphans' home remained normal or gained slightly in IQ, while those designated as "normal or superior-normal" showed a consistent loss. Children at all levels of ability represented who remained in an institution for the "feebleminded" were also found to show a consistent loss in IQ.

During the seven years following Wellman's initial report a series of eleven research papers were published by Wellman and her co-workers (Wellman, 1934, 1937, 1938a, 1938b; Coffey and Wellman, 1936; Crissey, 1937; Skeels and Filmore, 1937; Skeels, 1938; Skeels, et al, 1938; Skodak, 1939), all of which supported generally the view that the sort of environmental stimulation young children experience constitutes an important factor in their mental development.

> "The concept of intelligence as fixed and unmodified, or only modified within very narrow limits, is fast being replaced by a viewpoint which conceives of intelligence in functional terms. No longer can we say that an early intelligence test rigidly classifies a child for life in terms of mental potentiality" (Crissey, 1937, p. 7).

The Controversy. But the well ingrained idea that intelligence, as represented by the IQ, is an inherited, and therefore a "fixed" quality which characterizes the individual for life was not so easily to be "replaced" by such a radically different viewpoint. Certain students of child development immediately began to look for "erroneous assumptions," flaws in methodology and statistics, and "unwarranted conclusions" in the Iowa findings. Articles extremely critical of Wellman's reports began to appear in the psychological journals. One of the most violent of these attacks was a paper entitled, "The Wandering IQ: Is It Time to Settle Down"? by B. R. Simpson (1939). He assumed the role of one who is assigned the bitter task of "exposing" the devious operations of a colleague gone astray as he wrote:

> But to claim miracle working at Iowa that cannot be duplicated in other parts of the country in nursery schools, elementary schools, or high schools, that is much worse than nonsense and ought to be exposed as such, thankless as the task of exposing it may be (p. 366).

In general Simpson's attack apparently was based more on an emotional reaction to Wellman's point of view than upon a rational and objective evaluation of all of the evidence presented. Among his personal thrusts was a charge of "statistical incompetence under the influence of wishful thinking."

Another rather strong reaction was voiced by Florence Goodenough, certainly a highly respected "authority" in the field of intelligence and its evaluation. It should be mentioned at this point that in 1928, four years prior to Wellman's initial report, an article by Goodenough appeared in which she reported an "apparent increase" in average IQ in both of two matched groups, one of which had had a year's experience in a nursery school, while the other had not. The results of the study, however, lead to the conclusion that the "changes are largely, if not entirely, attributable to irregular standardization of the scale at the early ages and consequently should not be regarded as an indication of actual increase in intelligence" (Goodenough, 1928a, p. 368). No relationship between gain in IQ and length of attendance at the nursery school was indicated. In view of the findings of this early study, it is not surprising that Goodenough was highly critical of the Wellman papers.

In 1939 Goodenough published a paper entitled, "A Critique of Recent Experiments on Raising the IQ" (1939, pp. 73-79) in which she maintained that Wellman's conclusions, to a large degree, may be attributed to her acceptance of two basic hypotheses, "neither of which was accepted by most reputable scientists" of the day. These were:

(1) Intelligence tests given to infants and to children of two or three years of age have essentially the same predictive value for later mental development as those given after school age.

(2) The reliability of such tests for the measurement of mental status at the time they are taken is so nearly perfect that the effect of statistical regression due to errors of measurement may be safely ignored.

Dr. Goodenough continued to point out that infant tests had been shown to have no predictive value, and that due to their basic unreliability, a "regressive shift" takes place in successive testing of the same groups of children. This means that those who originally tested highest will appear as a group to lose in IQ because many of them by chance obtained IQ's that were too high, and that those who originally tested at the low end of the scale will, as a group, score higher at the next testing because many of their original scores were, by chance too low. She was also concerned with an apparent lack of controls over the possibility of bias on the part of the examiner in the Iowa test results. In general, Goodenough felt that little faith should be put in the results of the studies because of the lack of sufficient substantiating data.

During the following year Goodenough and Maurer (1940) published a report of a study in which nursery school children and non-nursery school children were compared as to change in IQ level after one, two and three years of nursery school experience on the part of the nursery school group. Interestingly enough their results were quite different from those of the Iowa studies. At the end of one year of nursery school experience the nursery school group and the non-nursery school group showed precisely the same average gain in IQ -- 4.6 points. Their general statement was that:

> None of the analyses that we have been able to make warrant the conclusion that attendance at nursery school has any measurable effect whatever upon the mental development of children. Those who have this training do no better on standardized intelligence tests than those who have not had it; they are neither more nor less advanced in school, and those who have attended longest and most regularly do not excel those whose period of enrollment was short and broken by frequent absences (p. 176).

This 1940 paper thus supported the idea of an hereditarily fixed potentiality as well as predetermined development, allowing no place for experience as having "any measurable effect whatever upon the mental development of children."

These conceptions about the nature of intelligence were even more strongly emphasized in Goodenough's review, also published in 1940, of "New Evidence on Environmental Influence on Intelligence," the purpose of which was to answer the question, "What effect, if any, is a specific environmental change likely to have upon the intellectual development of a particular group of children"?

In reviewing the Iowa studies Goodenough "accounted" for the differential changes which their data showed in terms of such factors as inadequate control groups, or control children that were not strictly comparable with respect to age, level of intelligence and test experience, the failure to take into account, and deal with "regressive effects," and other "indefensible" computational and statistical procedures. She concluded that, as a general outcome of their studies,

> The Iowa authors have developed an ingenious theory to account for this differential change, which, in spite of the extensive literature on the subject in the field of mathematical statistics, they continue to look upon as "real." The theory, as set forth in a number of the Iowa papers, appears to be that in any situation there is a critical point toward which the stimulating value of the environment tends to impel the intellectual development of the children. Those above this point lose for lack of further stimulation; those below it gain. The theory is an interesting one, but before it can be regarded as worthy of serious practical attention, it must be shown that the amount of the differential shift is greater than that necessitated by regression due to errors of measurement (Goodenough, 1940, pp. 313-314).

After reviewing a number of other studies of this general problem (Starkweather and Roberts, 1940; Kawin and Hoefer, 1931; Peterson, 1937) Goodenough reached two general conclusions as follows:

> 1. It is unsafe to assume that attendance at any unspecified nursery school is likely to bring about improvement in the mental ability of the average child, for in most investigations of this matter, no evidence whatever of such an effect has been found. It appears, therefore that ability to bring about intellectual improvement by this means is at best restricted to a few schools. The precise nature of the difference in the educational regimes of the schools that do, and of those that do not, claim to achieve these results has thus far not been made clear.
>
> 2. Analysis of the experimental studies dealing with this problem reveals many possible sources of error in the studies purporting to show a positive effect of nursery school training. Generally speaking, these studies have failed to maintain adequate control of basic variables. Their statistical techniques are also frequently of questionable validity, and in a number of instances, are certainly erroneous (Goodenough, 1940, p. 321).

The particular studies from Iowa which were perhaps most severely criticized because of extreme statements unjustified by the data, were their studies of foster children. These reports were largely based upon data on the later development of a group of 154 children who were placed in adoptive homes before they were six months of age, and a second group, 65 in number, adopted between the ages of two and five and a half. Certain statements which appeared in these studies were extreme, almost to the point of denying that the biological factor has any bearing at all upon later mental development. One such statement was:

> ....if there is an hereditary constitutional factor which sets the limits of mental development, these limits are extremely broad. Within these, environmental factors can operate to produce changes which for ordinary purposes may represent a shift from one extreme to another of the present distribution of intelligence among children (Skodak, 1939, pp. 131-132).

Another statement from the same report based upon the study of only 16 illegitimate children, is equally controversial:

> ....the mental development of children of feebleminded mothers and the most inferior true-family backgrounds is indistinguishable from that of children whose mothers are not feebleminded (Skodak, 1939, p. 104).

The controversy aroused by the Iowa studies naturally stimulated work in this field by a number of other researchers. In a study by Frandsen and Barlow (1940) a small average gain in IQ after experience in a "special nursery school" was reported for their experimental group, but this gain was not significantly greater than that made by their matched control group. The children in this study, 30 in the experimental group and 28 in the control group, were two to five years of age.

McHugh (1943) published a monograph dealing with preschool experience in relation to IQ changes in a group of 91 children, mean age of 62 months. After an initial Binet test they were given thirty three-hour sessions of preschool experience, after which they were again tested. Among McHugh's conclusions were:

> 1. Children do make significant mean gains in IQ score as a result of such preschool experience.
>
> 2. IQ gains from preschool experience are "adjustment gains" rather than growth in intelligence.

Peters and McElwee (1944) made a somewhat different approach to the problem. They first "analyzed intelligence into its components," then they set up these components as specific educational objectives for the nursery school experience. Their purpose was then to measure the outcome of this "highly analytic and purposeful procedure." The children of this study were only six in number, ranging in age from 2 years 9 months to 4 years 2 months. Their IQ's were below 100 and they were rated low in "initiative" and in "quickness to learn" by their teachers. They were from environmentally deprived homes.

At the beginning and at the end of the eight month experimental period, the Stanford-Binet, the Vineland Social Maturity Scale and the Merrill-Palmer Personality Rating Schedule on "sociability" were administered to them. The results were as follows:

> The average IQ gain was seven points which approaches, but does not reach statistical reliability.
>
> The mean gain in social quotient was 42 points, which even with such a small group is highly significant.
>
> The gain in mean rating on sociability was also significant (20 points).

This study, of course, is open to a number of criticisms. It is likely, for instance, had there been a control group, that the small gain in mean IQ would have had no significance.

Another different approach to the study of environmental influences was made by Bradway (1945). A group of preschool children selected as representative of the general population had been examined with the Stanford-Binet scale. Ten years later and with no contacts maintained in the interim, they were re-examined with the same scale. Several environmental factors in the lives of the children who had shown the most significant changes were examined and evaluated. Environmental information was obtained by interviewing the mothers and the children without knowledge on the part of the interviewer as to whether the child's IQ had shown an increase or a decrease. The results are summarized as follows:

1. The mean IQ of the increase group had increased 9 points, while the mean of the decrease group had decreased 17 points.

2. At the initial testing the mean IQ's of the two groups were within 6 points of each other. Ten years later the difference between their mean IQ's was 30 points.

As to the factors related to these changes, the findings were that:

3. The mean vocabulary score of the mothers of the increase group were significantly higher than the mean score of the mothers of the decrease group.

4. The mid-parent intelligence scores of the parents of the two groups also differed significantly.

The general conclusion was that the IQ changes which took place between preschool and junior high school were related to environmental factors, but the question as to whether the particular factors considered in the study "operated primarily as environmental stimuli or whether they were for the most part merely related to inheritable intelligence which had not been fully realized at the time of the initial examination" could not be answered from the data.

In general the studies so far examined which have concerned themselves with the evaluation of specific environmental factors in relation to mental development have resulted in inconsistent and contradictory findings. Attempts to isolate and to specify particular effects of experience in home reared children have met with little success. On the other hand, studies of infants reared in orphanages in comparison with home-reared or "mother" reared children have been rather consistent in showing wide differences in rate and level of behavioral development favoring the latter.

A feature held in common by the particular orphanages studied was a minimum of environmental variation for their infant inmates. In a study by Rene Spitz (1949) two groups of infants were compared, one reared in a "nursery" by their own mothers, the other in a foundling home by an over-worked nursing staff, one nurse having to care for 8 to 12 babies. The important difference in the two situations according to Spitz was the amount of environmental variation and stimulation, and especially conditions allowing for "emotional interchange" between the infants and those caring for them. Comparisons between the two groups were drawn in terms of developmental quotients (D.Q.) representing "six sectors of personality" i.e., level of perception, mastery of bodily functions, social relations, memory and imitation, manipulative ability, and intelligence. The contrasts between the two groups over a 12 month period were striking. At age 2 to 3 months the average D.Q. of the mother-reared children was 95 as compared with an average of 130 for the foundling home children. At ages 8 to 10 months the average D.Q. of the "nursery" children had risen to 110 while that of the foundling home group had dropped to around 72. In summary Spitz wrote,

> While the children in "Nursery" developed into normal healthy toddlers, a two-year observation of "Foundling home" showed that the emotionally starved never learned to speak, to walk, to feed themselves. With one or two exceptions in a total of 91 children, those who survived were human wrecks who behaved either in a manner of agitated or of apathetic idiots (Spitz, 1949, p. 149).

Spitz pointed out that in both institutions the infants received adequate food, the housing was excellent and medical care was equally adequate.

Similar findings were reported by Goldfarb (1943). He compared the development of children who were placed in an institution during early infancy and remained here until about three years of age and then placed in foster homes, with other children who were placed immediately in foster homes. The two groups were from much the same hereditary background. His conclusion was that prolonged residence in the institution during infancy was profoundly detrimental to the children's psychological development. For the institution, and foster home children respectively the percentages showing defective speech development were 80 and 15. Comparable percentages showing mental retardation were 37.5 and 7.5, and those showing educational difficulties were 42.5 and 15.0.

The findings of Spitz and Goldfarb, quite conclusively demonstrate the retarding effects of a lack of stimulation and environmental variation upon the behavioral development of young infants. Such findings, however, are open to different interpretations relative to the issue as to whether maturation is the sole factor in the development of mental capacity, or whether environment also is a factor of importance. Goldfarb, Spitz, Bowlby (1951), and others assume that low ability to perform in an intelligence test, or "defective speech development," on the part of institution children is indication of retarded development of mental capacity. In terms of this assumption, experience in adjusting to environmental variation - in coping with an ever changing situation - is essential to the development of intellectual potentiality.

Dennis and Najarian (1960), however, disagree with the above interpretation. They studied the development of children reared in a Lebanese foundling home, in comparison with children brought to the Well Baby Clinic of the American University Hospital of Beirut, Lebanon. Their objective findings were similar to those of Spitz. The average "DQ" for the institution babies over the age period from 3 to 12 months was only 63 while that of the Well Baby Clinic group was 101, a highly significant difference. In this age range all of the clinic group tested above the mean of the foundling home group. Not one of the latter group between the ages of 3 and 12 months had a DQ above 95. In discussing the similarity between their findings and those of Spitz, Dennis and Najarian wrote:

> Spitz's data and ours agree in finding that environmental conditions can depress infant test scores after the second month of life. We disagree with Spitz in regard to the interpretation of the cause of the decline. We believe that Spitz's data, as well as ours are satisfactorily interpreted in terms of restricted learning opportunities. We suggest that an analysis of the relationship between test items and the conditions prevailing in the Foundling Home would reveal that retardation could readily be explained in terms of restriction of learning opportunities (Haimowitz & Haimowitz, 1960, p. 187).

Dennis and Najarian, then, believe that the unvarying institution environment does not necessarily interfere with the actual developmental "unfoldment" of mental potential, but simply hampers learning; that life in the institution does not provide for the infant the opportunity to learn to perform the tasks, or to respond to the kinds of stimuli presented to it in the test situation, to nearly the same degree as does the ordinary home environment. From this point of view, the child may have possessed this capacity to learn these particular tricks, but did not have the opportunity to learn them. This intellectual capacity to learn had been developing all along in spite of stimulus deprivation. Thus, in terms of this theory, the environmental factor is not important in the development of intelligence defined as capacity to learn.

If this latter interpretation were correct - if the retarding effects of a depriving environment were simply the absence of learned responses due to the lack of opportunity to learn, rather than a retardation in the development of the capacity to learn - than all that would be needed to correct the condition in a given child would be to place him in an environment rich in opportunity to learn. Further findings by Goldfarb (1947) strongly suggest that this is not the case. In his study of the adjustments of a group of adolescents who had spent the early months and years of their lives in the institution, Goldfarb found the ill effects for the early environment were still with them. In discussing a particular case, he concluded that,

> ....even after years in a foster home he still demonstrated isolated, infantile impoverished character trends. His intellectual growth potential has been equally frustrated and he reacts to problems on a highly concretistic level that verifies the absence of an intelligent approach to new adjustments (p. 449).
>
> The investigation was begun in the hope and belief that an understanding of the factors associated with those individual variations in adjustment might help in the formation of a treatment program to counteract the effects of the privation process. The present findings are not encouraging. We have been able to do relatively little once the damage has been wrought (p. 455).

In summary the evidence seems generally to support the conclusion that experience in coping with a stimulating environment is a factor of importance in mental development. The assumption that mentality is "fixed" and determined in the sense that it is somehow derived automatically from the somatic and neural structures as they develop through the "mechanism of maturation" is hardly tenable. As Baller and Charles (1961) puts it, "optimum development of both learning capacity and learning achievement demand personal involvement of the learner and interaction with a rich environment" (back of page 245).

The Interaction Point of View. In view of the evidence then, "fixed intelligence" in the sense that the rate and pattern of its development are predetermined by the genes is no longer tenable.

In another sense, however, intelligence *is* "fixed." The idea of a "fixed intelligence" in the sense that the individual is endowed with a particular intellectual *potentiality* which is set by the genes surely is tenable. The idea, of course, is distinct from the idea of predetermined development. There is really no inherent connection between these two concepts. There is no necessity to assume that intellectual development, within the limits set by the genetic factor is wholly a function of maturation. The degree to which this fixed upper limit is approached by a given individual may be largely determined by the "nurture" factor - by the nature of his environment and his encounters with it.

Obviously a child's capacity to function effectively in an ever changing environment increases with age. Intelligence tests from the beginning have been designed to "measure" these changes, and in terms of test results, curves of increase in capacity have been plotted. These curves, furthermore, show a leveling off and a gradual cessation of increase as the individual approaches adulthood. Clearly these are curves of change in *capacity* to learn or to perform adequately in new or difficult situations and, since they do "level off" with adulthood and thus parallel generally the curves of organismic development, they undoubtedly are related to, and reflect generally the course of maturation of the neuro-effector system. It does not necessarily follow, however, that the development of the capacity thus depicted is purely maturational in nature - that it somehow arises automatically from the "natural unfolding" of an inherently designed anatomical structure uninfluenced by experience. On the contrary there is increasing evidence that experience is always an important determining factor and that richness of environmental stimulation has much to do, in each case, with the extent to which mental development at any point in time has actually approached the level which the structure at its particular stage of development has made possible. The development of the intellect, like all development, comes about through the interaction of the organism (with its fixed developmental potential) and its environment.

As we have seen, this point of view was expressed by Harvey A. Carr back in 1926. In 1934, when the controversy over the effects of environmental change on mental development was at its height, Stoddard and Wellman proposed that,

> A theory of mental development for which there is a small amount of supporting evidence and nothing as yet contradictory in it, may be built up along five main lines.
>
> 1. That definite limits to mental development are set by heredity, in increasing order of specificity as we go from species to race, to family, and finally to the individual.
>
> 2. That inherited mental ability, if allowed to function below its response potentialities, decreases relatively (that is, the intelligence quotient declines).
>
> 3. That inherited mental ability, in steady combination with stimulation appropriate to its needs, has the appearance of relative increase (that is, the I.Q. increases).
>
> 4. That in certain ways, the nervous system transforms external patterns into adequate internal stimuli and drives whatever mental mechanism is present accordingly.
> The mental range, for a given inheritance, runs from flabbiness to athleticism.

> 5. That in present day life, there are strong forces definitely subversive to mental development in individuals, in that they are substitutes for thinking, or distractions designed to discourage the process (Stoddard and Wellman, pp. 178-179).

This interaction theory of Stoddard and Wellman undoubtedly would be more widely accepted today than it was in 1937 as an explanation for IQ flunctuations and changes.

The current rather widely accepted view of mental development as a process of interaction is quite clearly expressed by Escalona and Moriarty (1961):

> Certainly the authors are not alone with their assumption that intelligence must be viewed as an interaction phenomenon. With apologies to Piaget's and other developmental theories we shall crudely describe intelligence as the result of a continuous stream of transactions between the organism and the surrounding field, which field is the sum total of physical and social environmental conditions. Since the organism brings to each adaptive act certain properties, both those of intrinsic biological nature and those which are the result of the impact of previous experience upon the organism, it is correct to say that intelligence development is at all times dependent upon what the organism is like. Yet, since it requires environmental circumstances to mobilize the organism, and since the kind of transaction which develops depends on the objective content of that to which the organism must adapt, it is equally true to say that the development of intelligence depends at all times on the experience encountered by the growing child (pp. 598-599).

Perhaps the strongest recent expositor of interactionism in relation to the development of intelligence is J. McV. Hunt (1961). He marshalls the evidence in support of this point of view from research in learning, learning sets, and the effects of early experience, from the programming of electronic computers, and especially from the observations of Piaget on the development of logical thinking and intelligence in children. He presents an extensive and systematic review of Piaget's "experiments" and stresses the fact that a basic concept in all of Piaget's theorizing about the growth of intelligence was the concept of interaction between organism and environment involving the complementary processes of accommodation and assimilation.

> The more new things an infant has seen and the more new things he has heard, the more new things he is interested in seeing and hearing; and the more variation in reality he has coped with, the greater is his capacity for coping. Such relationships derive from the conception that change in circumstances is required during the early sensorimotor stages to force the accommodative modifications in schemata and the assimilation of these modifications that, in combination, constitute development" (Hunt, 1961, p. 262).

In his concluding chapter Hunt points out some serious implications of the current prevailing conceptions and convictions about the nature of intelligence and how it develops in children in relation to the care and education of children. He writes:

> In the light of these considerations it appears that the counsel from experts on child rearing during the third and much of the fourth decades of the twentieth century to let children be while they grow and to avoid excessive stimulation was highly unfortunate. It was suggested in the text above that perhaps the negative correlations found between intelligence test scores for the first two years and the late adolescent level of intelligence may possibly be attributable to such counsel, inasmuch as it would be those educated people at the higher levels of tested intelligence who read and can act in terms of what they read who would have been most likely to follow this advice. The problem for the management of child development is to find out how to govern the encounters that children have with their environments to foster both an optimally rapid rate of intellectual development and a satisfying life (Hunt, 1961, pp. 362-363).

Points of View Regarding the "Structure" of Intelligence

Another important area concerning which theoretical positions have varied and changed over the years is the question of the structure or composition of intelligence. As we have already noted, the earliest attempts to measure intelligence were made in terms of separate "Faculties," processes or abilities. Out of these early attempts at measurement, and those which followed, however, there gradually developed the concept of "general intelligence," of which the various functions tested were expressions.

The Concept of General Intelligence. Binet in his writing gave no single, well formulated definition of intelligence but he did offer a number of descriptions of it, each stressing particular aspects of, or functions involved in mental activity. It is clear, however, that underlying these descriptions was the idea of a general intelligence, complex but unified. Intelligence, he wrote, is compounded of "judgment, common sense, initiative, the ability to adapt oneself." "To judge well, understand well, reason well -- these are the essentials of intelligence" (Binet and Simon, 1916, p. 42). William Stern (1914) on the other hand, offered a definition which clearly differentiates "general intelligence" from special talents and abilities, and from knowledge and mere information. "Intelligence," he said, "is a general capacity of an individual consciously to adjust his thinking to new requirements. It is general mental adaptability to new problems and conditions of life."

Charles Spearman (1904,1914, 1927) went farther than merely to offer a definition. Through the application of correlational, and other statistical techniques, he developed a theory known as the two-factor theory of intelligence. The first of these factors he referred to as "g," or a general factor identified as general intelligence. He conceived of g as a "general fund of mental energy." In summarizing his discussion of the nature of g he wrote:

> ....the facts of general psychology - quite apart from those of individual differences - strongly support the suggestion of mental energy and engines. Moreover, such an energy would seem to be just what is wanted to explain g (Spearman, 1927, p. 135).

In his thinking this general factor is involved in each specific function or performance, although the second factor is the individual's "specific capacity for that particular kind of performance."

Spearman's theory stimulated much interest and discussion. Other workers using Spearman's own formulae often were unable to duplicate his findings and thus came to challenge his theory. In opposition to his concept of a general factor was the contemporarily widely accepted view that intelligence consists of many specific abilities not bound together by a common factor (Thorndike, 1914). Spearman's two-factor theory was not generally accepted but the concept of general intelligence has predominated through the years. Some of the most widely used tests of intelligence, such as the Stanford revisions of the Binet scale, were designed and constructed in terms of the concept of general intelligence. The assumption was that all types of mental activity are very largely functions of a single general factor, and that it matters little what particular combination of mental tasks or performance is called for in an intelligence test scale, since they are all related to, and therefore tap, the same general factor. At least in part as a result of this view of the structure of intelligence, many of the scales currently in wide use, show little uniformity of content one with another; neither do they show any internal uniformity from age level to age level.

Concepts of Multiple Mental Abilities. As indicated above the alternative view of the structure of intelligence was the concept of specific abilities. Thorndike (1914), an eminent proponent of this view, wrote that "the mind must be regarded not as a functional unit, nor even as a collection of a few general faculties which work irrespective of particular material, but rather as a multitude of functions each of which involves content as well as form, and so is related closely to only a few of its fellows, to the others with greater and greater degrees of remoteness" (p. 366).

Spearman's two-factor theory and his famous tetrad differences technique for demonstrating the presence of a two-factor pattern in a set of correlations, as we have already noted, did give substantial statistical support to the theory of general intelligence. In terms of Spearman's findings, every test item in a scale was thought to be composed of the g factor -- the universal --, plus a specific factor peculiar to it alone, but his work also stimulated a great deal of discussion and further research with the use of his technique. In certain of these studies it was found that the intercorrelations between some pairs of tests were too high to be accounted for by their relatively low loadings of g. In other words, they involved a common factor in addition to g. These additional common factors came to be known as "group factors." Spearman himself, and his coworkers, came to recognize a number of such group factors, namely, verbal ability, numerical ability, mental speed, mechanical ability, attention and imagination. Concurrently, other analytical procedures designed to bring out the group factors in a battery of tests, were being experimented with. As a result, a number of methods of multiple factor analysis have been invented, among the best known of which are those of Thurstone (1935), Hotelling (1933), and Kelley (1935). Until the advent of the electronic computer, perhaps the most widely used of these was Thurstone's centroid method of multiple factor analysis.

In 1938 Thurstone reported a study in which he applied his own technique to the analysis of the intercorrelations among 57 tests designed to measure general intelligence. The data were the scores of 240 university students. Without benefit of the modern computer, this was a computational feat of no mean proportions.

The analysis was carried to 13 factors. The first six factors accounted for the major portion of the variance. Factors

10, 11, 12, and 13 were finally discarded as having no significant psychological meaning. Thurstone's tentative designations of the nine meaningful factors were Spacial, Perceptual, Numerical, Verbal Relations, Memory, Word Fluency, Induction, Restriction and Deduction.

In a later study, Thurstone and Thurstone (1941) extended their search for "primary mental abilities" downward to younger ages. On the basis of this study the Thurstones concluded that the following six factors met their criteria of primary abilities:

V	Verbal-comprehension
W	Word-fluency
S	Space
N	Number
M	Memory
I	Induction

In other words, "general intelligence" was found to consist of at least six mental abilities. However,

> "none of these factors is considered as fixed, indivisible, or non-combining; each one is dependent for its validity on the nature of the tests and the population examined. But each primary ability behaves as a functional unit that is strongly present in some tests and almost completely absent in many others" (Stoddard, 1943, p. 165).

A Three-Dimensional Model of Intellect. Two other investigators of note in the use of multiple factor analysis as applied to the problem of discovering and identifying the components of intelligence are Cattell (1953) and Guilford (1956, 1960). Guilford and his co-workers especially, in connection with their "Aptitudes

Project," at the University of Southern California, sponsored and financed by the U. S. Office of Naval Research, have concentrated particularly on the study of cognitive and thinking abilities.

> The results from the Aptitudes Project that have gained perhaps the most attention have pertained to creative-thinking abilities. These are mostly novel findings. But to me, the most significant outcome has been the development of a unified theory of human intellect, which organizes the known, unique or primary intellectual abilities into a single system called the "structure of intellect" (Guilford, 1959, p. 469).

This model makes room for a large number of factors, each representing a distinct ability. These factors, however, can be identified and ordered in terms of a three-dimensional system of classification. In other words, the model consists of three sets of classification categories. One of these bases of classification is in terms of mental operations. Guilford has identified five categories of operations as follows: cognitive processes, memory, convergent thinking, divergent thinking, and evaluation. These categories in Guilford's conceptualization, constitute the basic kinds of mental processes of "operations" of the intellect.

The second sort of classification is in terms of the kinds of content involved in the intellectual operations. The factors of intelligence so far identified involve the four following kinds of content: figural content, i.e., concrete material actually perceived through the senses; symbolic content composed of letters, digits, and other conventional signs, usually organized into general systems, such as the alphabet or the number system; semantic content in the form of verbal meanings or ideas; and *behavioral* content representing the area of inter-personal relationships sometimes called "social intelligence."

The third classification is in terms of the outcomes or products of the various "operations" as they are applied to any of the three kinds of content. There are six kinds of products. In the words of Guilford (1959),

> When a certain operation is applied to a certain kind of content, as many as six general kinds of products may be involved. There is enough evidence available to suggest that, regardless of the combinations of operations and content, the same six kinds of products may be found associated. The six kinds of products are: units, classes, relations, systems, transformations, and implications. So far as we have determined from factor analysis, these are the only fundamental kinds of products that we can know. As such, they may serve as basic classes into which one might fit all kinds of information psychologically (p. 470).

Guilford's theoretical model of the intellect in terms of its "structure," is thus represented as a solid three-dimensional figure, the dimensions being the three sets of classification categories - the operations, the contents, and the products. This conceptualization allows for 120 separate factors of intelligence, many of which have already been identified and measured. By dividing the model into vertical layers, or slices, in terms of the five categories of operations, we have five sets of 24 small cubes each. Each of these five sets, then, represents a particular category of mental abilities. Thus there are the cognitive abilities, the memory abilities, the divergent thinking abilities, the convergent thinking abilities, and the evaluative abilities. According to the model, each of the 24 cognitive abilities, for example, is an ability to deal with a particular kind of material, such as "symbolic content," for example. The outcome or "product" of that particular operation would then be the "cognition" of a symbolic unit," or a symbolic system, etc.

Since the early days of mental testing, then, we have seen an interesting shift in the nature of the prevailing conceptualizations of the "structure" of intelligence. As techniques of data analysis have developed, the trend in conceptualization has been away from the original view of intelligence as a unitary general ability presumed to account for individual differences in any and all sorts of mental functioning, toward greater and greater complexity of conceptualization encompassing a multiplicity of factors. First, in addition to g, certain specific, or group factors were postulated to account for the observed correlations among test variables. More refined analysis, however, soon led to the notion of multiple factors, as g became broken down into a relatively small number of "primary mental abilities." The advent of the electronic computer, of course, has led to a tremendous increase in analytic capacity in terms of the sheer number of test variables that can be handled in a single analysis, and in terms of speed of computational operations. This has led to the concept of a great variety of relatively independent mental abilities. The job of extracting these factors, identifying them and testing them is under way. The primary concern of the present project, however, is the problem of the assessment of mental abilities during infancy and early childhood.

The Problem of Developmental Change in Mentality

It has been shown repeatedly (Furfey and Muehlenbein, 1932; L. D. Anderson, 1939; Bayley, 1933, 1955; John Anderson, 1940; Cattell, 1947; Cavanaugh, et. al., 1957) that mental tests administered during the first year of the child's life are of no practical value in predicting later tested intelligence. Furfey and Muehlenbein (1932) for example set out specifically to determine the predictive value of the Linfert-Hierholzer scale administered during the first 12 months. They found no significant relationship between the infant test scores and Stanford-Binet scores obtained four years later. They concluded that since the other published infant scales were similar to the Linfert-Hierholzer in content and general make-up, their results called into question the predictive value of all infant mental tests.

A study by L. D. Anderson (1939) produced similar results with infant tests constructed by the "usual percentage-of-success method." Zero, or insignificant relationships were found between scores at 3, 6, 9, or 12 months, and Stanford-Binet scores obtained at age 5 years.

More recently Bayley (1955) obtained correlations of near zero between scores on the California First Year Mental Scale administered earlier than one year of age, and retest scores at 18 months. With retests at later preschool ages the correlations were even negative, as high as -.21. Predictive results with the Cattell Infant Test were found to be very little better (Psyche Cattell, 1947).

Correlations between IQ's at 3, 6, and 9 months, and later 3-year IQ's were .10, .34 and .18.

Even though these tests have been found to be valuable aids in the over-all appraisal of health and developmental status of babies, this lack of predictive value has raised serious doubts as to their validity as tests of intelligence. John Anderson, for example, was expressing his doubt when he states that,

> ....the adequacy with which a particular test measures what it purports to measure has to be determined in terms of its correlation with tests at later ages... In making the best possible prediction of terminal status, we will also make the best measurement of present status in so far as our concern is with potentiality rather than achievement..
>
> Infant tests as at present constituted measure very little if at all, the function, which is called intelligence at later ages (1940, p. 401).

Anderson thus made it clear that he could not accept an appraisal of an infant's mental status as valid unless it reliably fore-told future or terminal mental status. It is interesting to note in this connection that the validity of the criterion (usually the Stanford-Binet scale) in terms of which the infant tests are judged, was generally taken for granted without question, even though in some instances the criterion test had been less adequately standardized than the infant test in question (Kawin, 1934).

Because of this concern about predictive validity, some otherwise rather promising tests of infant abilities may have been abandoned. The Linfert-Hierholzer scale, for example, carefully standardized on 300 infants under one year of age, some 50 of which were tested repeatedly at different ages, has never been used to any extent because it was found not to predict Binet IQ.

Quite obviously, a reliable means of appraising intellectual potential in infants and thus to be able to predict future academic ability would be of considerable practical value, particularly to adoption agencies and prospective adoptive parents. The inadequacy of presently available tests for making such early predictions is not only a matter of practical concern but constitutes a theoretical problem of considerable import. It raises the question as to the nature of infantile mentality and how it differs from intelligence in later childhood. There have been in the past two principle points of view regarding this difference and the nature of mental development. A brief consideration of each of these viewpoints follows.

The Assumption of Quantitative Change. The traditional view is that intelligence is fixed, that it develops naturally with the maturation of the nervous system, and therefore, that the IQ, when reliably determined, represents a stable characteristic of the individual. This view clearly includes the idea that there is no qualitative change in the nature of intelligence during development, but rather that the intellect develops through steady and continuous quantitative increments from birth.

Although this viewpoint had its heyday in the early 1930's it is difficult, nevertheless to find definite statements by the psychological writers of that period regarding the nature of infantile mentality or the nature of developmental change in intelligence. The essentially quantitative nature of mental growth was generally assumed without thought of discussion, or of a consideration of other possible alternative views. The whole intelligence testing movement, of course, was based upon this assumption of quantitative change. Inherent in the very concept of measurement is the implication of a scale or system of quantitative units. Thus if intelligence was to be measured, it must be measured in quantitative terms.

Binet on Mental Developmental Change. It is quite clear that Binet approached the task of measuring individual differences in mental capacity among children with the assumption that these differences were essentially quantitative in nature.

As was mentioned earlier, although he seemed generally to regard intelligence as inherited, Binet was opposed to the idea of intelligence *per se* as a "fixed quantity." Obviously when he made that statement he was not applying the term to a genetically determined potentiality, but rather to a capacity which changes with age. And he was quite specific as to the nature of this changing capacity.

> It seems to us that in intelligence there is a fundamental faculty, the alternation or the lack of which is of the utmost importance for practical life. The faculty is judgment, otherwise called sense... A person may be a moron or an imbecile if he is lacking in judgment, but with good judgment he can never be either. Indeed the rest of the intellectual faculties seem of little importance in comparison with judgment (Kite, 1916, pp. 42-43).*

---

* Quotations from the writings of Binet are taken from Elizabeth S. Kite's "The Development of Intelligence in Children," translated from the writings of Binet and Simon in *L'Annee Psychologique*, 1905, 1908.

Binet, at another point, was more specific as to what he meant by "the rest of the intellectual faculties." He indicated that he regarded such processes as "sensation and perception" as "phenomena of intelligence" and as "intellectual manifestations," and not essential constituents of intelligence as such. He saw these functions as "distinct from, and independent of judgment" (Kite, 1916, p. 120).

As to the nature of the change in the capacity to judge (intelligence) Binet undoubtedly regarded it as largely quantitative. "Judgment," from his point of view, is a "faculty" which grows in a quantitative sense from perhaps a zero amount at birth to varying degrees as children develop.

> It (the development of intelligence) consists in part in the increase of the faculty of comprehension and of judging, at least this is probable. A child understands less and judges with less penetration than an adult; it consists also in the increase of acquisitions of every sort. But these are perhaps secondary characteristics which one may lack without compromising his maturity (ibid. p. 257).

On the other hand, although he stressed the quantitative aspect of developmental change, Binet apparently saw in mental development more than the simple addition of quantitative increments. "The child differs from adults," he wrote, "not only in degree and quantity of his intelligence, but also in its form. What this childish form of intelligence is, we do not yet know" (ibid, p. 183).

Genetically Endowed Potentiality. Writers following Binet quite obviously were primarily concerned with the measurement of

2

an assumed general quantity called intelligence. Psychologists became preoccupied mainly with the construction and standardization of tests and with the quantitative measurement of growth changes and individual differences. They gave relatively little space in their writings to a consideration of the essential nature of the thing they were trying to appraise, or with the nature of its developmental changes. There was, as we have seen, considerable variation in the usage of the term intelligence. But in general however, there was an implicit, if not stated, assumption that the individual is endowed at conception with a fixed mental developmental potentiality which _is_ his intelligence, and that mental growth therefore is an orderly increase in capacity to function within the limits, and toward the full realization of that developmental potentiality which, _in itself_, does not change.

As we have already noted, the term intelligence is often used also to refer specifically to this growing _capacity_. In that sense intelligence changes in a quantitative sense but remains qualitatively the same. The child as he matures simply gains more of the capacity to sense, to learn, to solve problems, to profit from experience or to "judge," but always within the limits of his genetically determined potentiality.

Among those who have been quite explicit in their identification with this point of view is Florence Goodenough. She wrote as follows:

> That children reared from infancy in environments where intellectual opportunity is not lacking and where incentive to intellectual achievement is high are likely to reach a higher level of achievement than others of equal original endowment from whom both opportunity and

> incentives are poor is conceded by practically all who have considered the matter. But this is not equivalent to saying that either opportunity or incentive will bring about changes that go beyond the limitations set by the germ plasm. "Men do not gather grapes from thorns nor figs from thistles" (Goodenough in Carmichael, 1946, pp. 485-486).

But the idea of a genetically determined limitation upon the development of mental capacity is by no means peculiar to modern psychologists. Plato, writing 2000 years ago in his *The Republic* discussed individual differences in mentality in terms of original endowment. "When you spoke of a nature gifted or not gifted in any respect," he inquired, "did you mean to say that one man will acquire a thing easily, another with difficulty?....Would not these be the sort of differences which would distinguish the man gifted by nature from the one who is ungifted"?

The concept of the "gifted child" and of the mentally defective or retarded child clearly carry the implication of a genetically determined potentiality, and that differences with respect to this general quality are quantitative in nature.

The Theory of Quantitative Change and the Lack of Predictive Value in Infant Tests. The newborn quite obviously is lacking in ability on his own to cope with his environment and during the months of early infancy he will have enough growing time to develop only a limited degree of functional capacity. From the standpoint of quantitative developmental theory, that capacity (intelligence) would differ only in amount from time to time.

What then, about the fact that the results of infant mental tests, and those of tests administered at later school ages do not correlate significantly? On the theory that there is no qualitative difference in intelligence at these age levels, the general conclusion of course has been that the available infant tests do not measure intelligence. On the other hand, in the opinion of many of their users, these tests are valid instruments of measurement and do provide an accurate appraisal of the baby's mental status. If this is true then intelligence must change in kind or quality as well as in degree.

There are, however, a number of possible alternative explanations for the lack of predictive value of the infant tests. One hypothesis takes account of the profound immaturity of the infant subject. Time is required to establish stimulus-response connections and to develop patterns of behavior which constitute adaptive ability. Since intelligence at any level can be tested only in terms of what the subject can do, there is, by virtue of sheer immaturity, an extremely limited basis for testing infant ability or capacity. Not only is the tester limited by what the infant can do, but particularly is he limited by what he can get the baby to do even within his repertory of possible acts. Resistance in the mental test situation of three-year-old children has been found to be an important factor affecting test results (Rust, 1931). Overcoming the resistance that was present at the first testing was found to result in increases in IQ's (Kuhlman-Binet) of approximately 50 percent of the children, the increases ranging from 15 to 35 IQ points. It is not unreasonable to expect that resistance of similar magnitude and variability occurs among young babies.

Level of maturity is a factor here in that the baby is relatively immune to verbal persuasion. In his immaturity also he is subject to changing bodily conditions and fatigue.

The extent to which infant test performances represent their actual ability to perform may also depend upon certain more stable factors in their personal make-up. Differences in original temperamental nature, e.g., in activity level, characteristic reactivity to stimulation, general alertness, etc., might make for wide differences in performance between two infants with the same capacity level. By the same token, two infants with quite widely different inherent potentialities and capacities may actually exhibit equivalent levels of performance on the test.

Rachel Stutsman (Ball) was fully cognizant of these factors during the process of developing the Merrill-Palmer Scale. In her discussion of the role personality factors play in the test situation she stated that,

> ....back of every individual response is a group of variables which we have not been able to control and of whose relative strength we are totally uncertain. Self-reliance, initiative, self-consciousness, and persistence are such variables. At present it is impossible for us to isolate the effects produced by these non-intellectual traits from those produced by the intellectual abilities. Even in the most rigidly controlled test situation we are, of course, testing the ability of the individual to do a certain piece of work, and though the results may be called a measure of the individual's intelligence they can certainly be analyzed into causative elements, many of which are not ordinarily implied in that concept (Ball, 1931, p. 241).

Another area in which individual babies differ fundamentally is in pattern of development. The various aspects of over-all development do not keep pace with each other. Each child has an uneven "developmental front" and the pattern of that front is different for each infant. Thus, general expectations of performance at different levels of infancy fit the developmental patterns of different infants with varying degrees of accuracy apart from the question of mental potential. These differences in developmental pattern undoubtedly contribute to the predictive inadequacy of infant tests. Nancy Bayley (1955) took cognizance of this factor of individual differences when she wrote:

> The neonate who is precocious in the developing of the simpler abilities, such as auditory acuity or pupilary reflexes, has an advantage in the slightly more complex behaviors, such as (say) turning toward a sound, or fixating an object held before his eyes. But these more complex acts also involve other functions, such as neuro-muscular coordinations, in which he may not be precocious. The bright one-month-old may be sufficiently slow in developing these later more complex functions so as to lose some or all of his earlier advantage (pp. 807-808).

Other possible causes for the lack of infant test predictivity have been proposed. Escalona (1948, 1955), for example, has suggested that the situation might be improved by interpreting test results and thus modifying the scores in terms of clinical observations made at the time of the test. The scores would then be based, not simply on the items passed, but also to some degree on clinical estimates of the child's potential. "The same formal test results," she wrote, "may be interpreted differently in the light of concomitant clinical observations, thus increasing the predictive value of the tests" (1948, p. 281).

It has also been suggested that the inconsistent results of infant tests is due to their item composition. They are said to lack suitable kinds of items. Catalano and McCarthy (1954), for example, studied the prelinguistic vocalizations of a group of infants at an average age of 13 months in relation to the results from an intelligence test administered after an average interval of 31 months. Although their correlations were discouragingly low, the investigators suggested that prelinguistic vocalization might be better predictors of later intelligence than the usual infant scale items.

The lack of predictive validity of infant tests might also be, in part, a function of the amount and quality of stimulation the infant, even very early has experienced. In terms of the maturation theory, two infants with equivalent mental potentialities would presumably have developed about equal capacities for mental test performance, but, due to differences in sensory and perceptual experience - experience in coping with environmental stimulation - may not have acquired equally well the response patterns necessary to perform adequately on the test (Dennis and Najarian, 1957).

In view of the above considerations, it may be concluded that the lack of correlation between infant test results and the results of tests administered at later ages, although damaging, does not necessarily negate the assumption of qualitative invariability in the development of intelligence. The lack of infant test predictiveness is further considered in connection with other concepts of mental development.

Development as Qualitative Change. As we have seen, intelligence is often broadly defined as "ability" or capacity to cope with, or adjust to circumstances of living. From this point of view one naturally thinks of mental development as the elaboration or enlargement of the child's repertory of adjustive behavior patterns. This increase in adaptability is seen as something more than the mere addition of increments to a basic quality or ability. The implication is rather that new qualities, new patterns of adaptation emerge as the child grows and interacts with his environment.

Biologically speaking, the child has been growing by virtue of interaction with his environment since the moment of conception. By the time of his birth he has already developed certain simple behavioral, as well as physiological patterns in terms of which he is able to adapt to the conditions of his new external environment, and without which he could not survive. In the sense, then, that intelligence is the capacity to make adaptive responses, the neonate possesses a measure of intelligence. Even with his immaturity and complete dependence at birth, he is equipped with a certain repertory of simple adaptive response patterns.

Generally, these adaptive patterns are "standard equipment." They are characteristic of the species at the neonate level. Recent research, furthermore, has led to an increasing acceptance of the view that the young infant is not incompetent. The neonate possesses capacities for sensory discrimination and precision of adaptive response that have only recently been elicited and described (Bronshstein, Antonova, Luppova, and Sytova, 1958; Kessen and Leulzendorff, 1963).

It is largely in relation to these early modes of adaptation that individuality is first expressed. Each infant is unique - an individual different from all others in the expression and over-all patterning of these adaptive behaviors. He displays this individuality in his reactions to stimulation, in the way he nurses, in his crying, in his general activity level, in his patterns of sleeping and waking. Thus neonatal "intelligence" expresses itself, to some degree, uniquely in the individuality of each infant. It is an expression of "mental" development that has previously taken place. In the words of Arnold Gesell (1940),

> ....these neonatal expressions of individuality are largely the end product of the primary mental growth which was accomplished in the long period of gestation. From a biological standpoint there are no sharp transitions in the continuum of mental growth. Even birth does not bring about a unique and abrupt transition, because *in utero* the fetus has already anticipated to a great degree the reactions of early neonatal life. He has been prepared; the very arrangement and relationships of his neurons have pointed to the future. In preliminary and provisional form these relationships were laid down by intrinsic patterning prior to and independent of actual experience, this preliminary and prospective kind of patterning is mental maturation. It operates not only *in utero* but throughout the whole cycle of mental growth (pp. 12-13).

Gesell's View of Developmental Change. Gesell described mind as a process. This process is the organizing, integrating, controlling function of the individual. Mental growth (growth of the mind) for him therefore is a process of change, of increase in the effectiveness of the mind. "It is a process of behavior patterning." It is more than the addition of quantitative increments of what is always there. "The child's mind," Gesell (1940) wrote, "does not grow by a simple linear extension. He has a persisting individuality, but his outlook on life and on himself transforms as he matures. He is not simply becoming more 'intelligent' in a narrow sense of this much misused term. He alters as he grows" (p. 12).

He further identifies mind with the functioning of the neuromuscular system and the other involved physical structures, and "mental growth" with the maturation of these structures.

> As the infant grows and as the child grows, these nerve cells become organized into patterns of responsiveness, or into reaction systems. These neuron patterns determine behavior. They are influenced by the constitution of the blood, by endoctrine hormones, and by electro-chemical regulators; but in a fundamental sense the patterning of the mind is inseparably connected with the microscopic and ultramicroscopic patterning of nerve cells (ibid, pp. 11-12).

In further emphasizing his view of the great and primary importance of predetermined biological maturation, and of the secondary and limited role played by the environment in mental development, Gesell later wrote:

> The action system develops as a unitary whole, within which individuation arises, from intrinsic morphogenetic forces. These individuations have a fore-reference; they anticipate subsequent adjustments to the environment with provisional and preparatory arrangements. The environment does not engender the arrangements. Experience does not create them. The neural mechanism of walking is laid down before the child can walk. This in turn has been anticipated by pre-neural organizations.
>
> The primary motor attitudes and the basic initiative in psychomotor attitudes are thus endogeneous. We apply the term maturation to these intrinsic and prospective aspects of ontogenetic patterning. Environmental inflects and specifies but it does not engender the progressions themselves (1945, p. 163).

Gesell does not claim for his "Developmental Schedule" that it is primarily a test of infant intelligence, but rather he regards it as a normative device for appraising developmental status in young children. In fact he has little use for the term "intelligence." However, because of his close identification of "mental growth" with the maturation of the organism and thus with general functional development, the implication is clear that his schedule is designed to be a measure of "mental growth." In his rare use of the term intelligence he limits its meaning to the "adaptive" functions of the individual. Thus in his discussion of "the four major fields of behavior" covered in his schedule, i.e., motor characteristics, adaptive behavior, language and personal-social behavior, he described the second of these fields, "adaptive behavior," as dealing largely with intelligence. "Adaptive behavior," he wrote, "has been described as 'a convenient category for those varied adjustments perceptual, orientational, manual and verbal which reflect the child's capacity to initiate new experiences and to profit by past experiences'" (1940, p. 108).

In summary, it is quite clear that Gesell regarded mental development as the functional aspect of biological maturation - a process of change which is qualitative, as well as quantitative in nature. A further implication inherent in his viewpoint is that this process, in which environment plays only an "infecting" and "specifying," rather than an "engendering" role, takes place under the control, and within the limits of a genetically determined developmental potential.

Piaget's Epigenetic Theory. Jean Piaget and his associates at Geneva, Switzerland were the first actually to formulate a theory of mental development in which the nature of developmental change as such was a primary concern. Piaget conceives intelligence in very dynamic terms. He defines it as a process of organization - as "an assimiliatory activity whose functional laws are laid down as early as organic life and whose successive structures serving it as organs are elaborated by interaction between itself and the external environment" (Margaret Cook, translation of Piaget, 1952, p. 359). Three essential elements in his theory are given in this definition. First is the idea that intelligence is not an entity, not a quantity, not a latent potentiality, but a dynamic process, an activity. Secondly, he indicates the nature of that process. It is a process of organization and adaptation. The third point is that the process - the operation of these "invariant functional laws" - results in many variable structures, or "schemata" which in turn become the "tools" for further interaction with the environment through assimilation and accommodation. Piaget further elaborates on these essential points"

> ....intelligence constitutes an organizing activity whose functioning extends that of the biological organization, while surpassing it due to the elaboration of new structures...if the sequential structures due to intellectual activity differ among themselves qualitatively, they always obey the same functional laws. In this respect, sensorimotor intelligence can be compared to reflective or rational intelligence and this comparison clarifies the analysis of the two extreme terms (1952, p. 407).

In discussing the principles of assimilation and accommodation in their broader application, Piaget wrote as follows:

> ....it can be said that the living being assimilates to himself the whole universe, at the same time that he accommodates himself to it... It is therefore permissible to conceive assimilation in a general sense as being the incorporation of any external reality whatever to one part or another, of the cycle of organization. In other words, every thing that answers a need of the organism is material for assimilation, the need even being the expression of assimilatory activity as such... The functions of relationship, independently even from psychic life which proceeds from them, are thus doubly the source of assimilation (ibid., pp. 407-408).

Even though for Piaget, intellectual developmental changes from birth to maturity do constitute a fixed sequence of stages, each characterized by its particular organization of schemata, and thus being qualitatively different from the one from which it emerges, yet in his view there is complete continuity in the course of this development. "Intelligence thus germinates in life itself" and the same invarient processes are at work from the beginning in the organization of new schemata on the basis of the older, more simple ones. He states that new schemata "whose appearance marks each stage, are always revealed as developing those of the preceding stage" (ibid. p. 384).

The concept of need is also basic to Piaget's thinking. The function of the behavior patterns which he calls "schemata" is the satisfaction of need. This means "that the behavior patterns are from the onset the function of the general organization of the living body. Every living being constitutes a totality which tends to conserve itself and consequently to assimilate to itself the external elements it needs" (ibid. p. 389).

As is indicated in Piaget's definition of intelligence quoted above, the schema is in reality a tool of the process of adaptation (assimilation and accommodation). A schema is a behavior pattern and thus conceived, it alone constitutes the organizing activity - the functions of assimilation and accommodation. Thus it is by virtue of the activity of the existing schemata that new schemata are evolved. A schema, then, being a pattern of behavior, is "applied to the diversities of the external environment" and thus becomes increasingly generalized through its application to the varied aspects of the environment.

Observed developmental changes in mentality for Piaget, then, are qualitative in nature. Each "stage" of mental development encompasses patterns of behavior (schemata) which are new and different, but which, as Piaget puts it, are developments of the old patterns. His theory is truly one of epigenisis. The earliest "organic" stage gives rise to the next stage by the progressive production of new parts, or schemata which were previously non-existent, not preformed, and not predetermined.

Piaget and Gesell both see mental development as qualitative change, but as to the basis, or explanation of that change, they are far apart. Gesell places great emphasis upon the "mechanism of maturation" which brings about these qualitative changes through the natural unfoldment of a predetermined developmental design. Piaget, on the other hand, recognizes the important role of the environment. Mental development, for him, is not predetermined, is not a product of biological maturation, but is rather a result of continuous interaction between the individual and his environment through assimilation and accommodation.

From this point of view the character and the variety of environmental stimulation looms as a matter of vital importance. In the very early weeks of life the infant is incapable of direct, voluntary action on his own initiative. He can simply respond to changes and variations of environmental stimulation. Thus if he is to develop functionally his environment must provide changes and variation in conditions for him to cope with. As Hunt (1961) puts it,

> The more new things an infant has seen and the more new things he has heard, the more new things he is interested in seeing and hearing; and the more variation in reality he has coped with, the greater is his capacity for coping. Such relationships derive from the conception that change in circumstances is required during the early sensorimotor stages to force the accommodative modifications in schemata and the assimilations that, in combination, constitute development (p. 262).

This emphasis on the importance of stimulus variation for early mental development is appearing more and more frequently in the literature of infant development. A stimulus rich, but carefully ordered environment is a prime condition for optimum mental progress in infants and preschool children in the view of a member of present-day writers and investigators (Fowler, 1962a, 1962b; Hunt, 1961; Moore and Anderson, 1960). Discussions of the role of environment in mental development are also leading to some reinterpretation of early studies (Dennis and Najarian, 1957) of the effects of "maternal deprivation" in institution-reared babies (Ribble, 1943; Goldfarb, 1945; Spitz, 1949; Bowlby, 1952).

Piaget's theory of a developmental sequence of "stages" in mental development is based primarily upon his carefully recorded observations of his own three children. From these "experiments," as we have seen, he presents a picture of continuous transformations in the organized "structures" of intelligence. This series of "organizations," which are the products, as well as the tools, of assimilation and accommodation, begins at birth. But at birth the organizations consist only of congenital "sensorimotor schemata," and these continue to characterize intelligence throughout the first 18 months or two years of the child's life. During this period, the sensorimotor period, Piaget distinguishes six stages.

The first stage is characterized by the exercise of the ready-made, congenital schemata such as sucking, vocalizing, listening, looking, comprehension, etc. In the very beginning, as indicated above, there is simply passive release of these schemata by stimulation. There is, then, during this stage, a gradual shift to an active "groping." "The subject does not remain passive but, on the contrary, manifests the behavior pattern emphasized by Jennings: He gropes and abandons himself to a series of 'trials and errors'. That is, according to Claparede (1917), the origin of intelligence" (Piaget, 1952, p. 397). The reflex thus becomes "consolidated and strengthened by virtue of its own functioning." In other words it is assimilated to the child's needs and functioning. At the same time it becomes adapted or "accommodated" to the realities of the situation through the groping, trial and error process.

The second stage, the stage of "primary circular reactions" is characterized by the progressive coordination and assimilation of the ready-made schemata to form motor habits and perceptions. At this stage the "hereditary adaptations are doubled at a given moment, by adaptations which are not innate to which they are subordinated little by little. In other words, the reflex processes are progressively integrated into cortical activity" (ibid. p. 47). Three main developments are recognized during this stage. (1) Variations in schemata appear as a variety of stimuli become assimilated to them. (2) A reciprocal coordination among the schemata takes place. Thus hand movements, to a degreee, become coordinated with sucking, and things looked at become something to be reached for and grasped. (3) Although the child is unable to respond to a "vanished object," repeated stimulation by objects leads to "perceptual recognition" of them. This second period, according to Piaget's observations, extends from about the end of the first month of life to about age five months.

Stage number three, the stage of "secondary circular reactions" extends to about 8 or 9 months of age. Here again there is functional continuity" with the earlier stages "thus establishing complete continuity between increasingly complex structures." As contrasted with the earlier stage, when the baby's actions are "centered on themselves" - when he grasps for the sake of grasping- his actions in this third stage become "centered on a result produced in the external environment, and the sole aim of the action is to maintain this result: furthermore, it is more complex, the means beginning to be differentiated from the end, at least after the event" (1952, p. 157). The child now begins to show some anticipation of the consequences of his own acts, the beginning of "intentionality."

The fourth stage of the sensorimotor period is characterized by the "coordination of the secondary schemata and their application to new situations" (1952, p. 210). It is during this stage (8-9 to 11-12 months of age) that "the first actually intelligent behavior patterns" appear. Schemata, which heretofore were detached, now may be coordinated by the child in a single, more complex act with the aim to attain an end which is not immediately within reach. He thus puts to work, with his intention, schemata coordinated for a new purpose. In this way the child begins to show behavior in which means are clearly differentiated from ends. In his new relationships with his environment the "accommodation" aspect of his adaptations is especially apparent, as he begins to differentiate the self from the non-self and is able to search to a limited degree for the "vanished object." He also shows evidence of an implicit conception of causality and he appears to foresee events which are independent of his own actions.

Stage four, although it evolves from, and is a further development of the earlier stage, is clearly marked by qualitatively different behavior features. However, as Piaget points out, the child is still especially limited in the effectiveness of his coping by two conditions:

> In the first place, in order to become adapted to the new circumstances in which he finds himself - that is to say, in order to remove the obstacle or discover the requisite intermediate - the child at the fourth stage limits himself to intercoordinating familiar schemata, except for differentiating them through progressive accommodation while adjusting them to each other. In the second place, and through that very fact, the relations which the child establishes between things still depend on completed schemata of which only the coordination is new; besides they do not lead to the elaboration of objects entirely independent of the action... In short, the fourth stage, in so far as it is defined by the commencement of the coordination of schemata, appears more as a phase of initiation or of gestation than as a period of realization or accomplishment (1952, pp. 263-264).

The fifth stage, approximately the second year of life, by contrast, "is primarily the stage of elaboration of the 'object'." It is characterized by the formation of new schemata which are established through a sort of experimentation, a search for something new. These new schemata come about, in other words, through activity directed with more of a purpose to seek novelty for its own sake, rather than simply through the practice of acts that produced chance results as in the earlier stages.

This stage Piaget called the stage of "tertiary circular reactions." This type of reaction, he says, is quite different:

> ....if it also arises by way of differentiation, from the secondary circular schemata, this differentiation no longer imposed by the environment but is, so to speak, accepted and even desired in itself...the child manifests an unexpected behavior pattern. He tries, through a sort of experimentation, to find out in which respect the object or the event is new. In other words, he will not only submit to but even provoke new results instead of being satisfied merely to reproduce them once they have been revealed fortuitously (1952, p. 266).

This type of experimentation is seen in the child's deliberately letting an object, which he is holding, fall to the floor, watching intently what happens to it. He shakes objects, knocks them, listens to the sounds they make, throws them and watches them bounce or roll. He is constantly "experimenting in order to see." In the process of experimentation there is a constant accommodation of patterns (schemata) to the situation, and assimilation of new schemata. This naturally leads to "the discovery of new means" based on the apprehension of new relationships.

Piaget refers to this type of mental activity as "inventive intelligence." In his observations, one of its earliest manifestations was what he called the "behavior pattern of the support." The pattern, based upon perceiving the relation between an object and whatever is supporting it, consists of grasping the underlying support of the object in order to draw it within reach, e.g., child grasps a cushion and draws it toward him in order to obtain a box placed upon it. A "new means" of obtaining a desired object is thus discovered and utilized. And, in the opinion of Piaget, such new means can be discovered only through this "experimenting-in-order-to-see" activity which is characteristic of stage five.

Much groping, or trial and error behavior is manifest in the child's experimental attempts to obtain desired objects during this stage. Such behavior leads to an appreciation in him of spacial, and causal relations and temporal sequences. This "drive" to active exploration and manipulation of objects makes evident the great importance during this stage of an environment rich in appropriate stimulation.

The sixth and last stage of the sensorimotor, or "practical" intelligence appears in Piaget's thinking to be, in a sense, a transitional phase between the practical, and the "systematic" or deductive levels of intelligence. In discussing this stage, Piaget takes pains once more to emphasize the fact that the beginning of a new stage "does not abolish in any way the behavior patterns of the preceding stages and that new behavior patterns are simply superimposed on the old ones" (1952, p. 331). The patterns previously described are merely "completed behavior patterns of a new type." Something new comes into the picture (qualitative change) without "abolishing" the old.

This new type of mental activity is "governed by awareness of relationships and no longer by empirical groping" (Claparede). Piaget points out that authoritative writers from different points of view agree that "there exists an essential moment in the development of intelligence: the moment when the awareness of relationships is sufficiently advanced to permit a reasoned pre-vision, that is to say, an invention operating by simple mental combination" (1952, p. 332). Unlike the previous stage in which the "new means" are discovered through groping (pulling the cushion toward one to obtain object resting on it), new means now may be "invented." The child foresees which acts will succeed and which will fail without empirically testing them.

The mental process is one of deduction. The resulting procedure is "new" in the sense that "it results from an original mental combination and not from a combination of movements actually executed at each stage of the operation" (1952, p. 341).

Along with invention, or rather one aspect of it, is "representation." Sensorimotor groping gives way to the representation, ideationally, of the various possible acts and how they must be combined to achieve the desired end. "To invent is to combine mental, that is to say representative, schemata and, in order to become mental, the sensorimotor schemata must be capable of intercombining in every way, that is to say, of being able to give rise to true inventions" (1952, p. 341).

Throughout this whole sequence of early epigenetic stages portrayed by Piaget, the factor of experience plays an extremely important role. Coping with the environment, that is, interaction between the individual and his environment (assimilation and accommodation), constitutes experience. "At every level, experience is necessary to the development of intelligence" (ibid. p. 362).

Other Conceptions of Intelligence Which Imply Qualitative Change. It is quite clear from the foregoing discussions that any statement about the nature of developmental change in intelligence must be formulated in view of what specifically is meant by the term "intelligence." One widely accepted view of the development of behavior in general, at least in its early stages, is that it consists of the differentiation of specific functions and patterns out of "mass," undifferentiated, uncoordinated activity. In the words of Jensen (1932), "stimulation of almost any group of receptors by almost any kind of stimulation will lead to a response in almost any part of the organism that is set to respond." As the baby develops, movements become specialized and integrated into patterns resulting in the differentiation of specific functional patterns of behavior out of the generalized, undifferentiated activity. Garrett (1946) has advanced the hypothesis that the development of intelligence is likewise a process of differentiation.

For Garrett, intelligence,

> includes at least the abilities demanded in the solution of problems which require the comprehension and use of symbols. By symbols I mean words, numbers, diagrams, equations, formulas which represent ideas and relationships ranging from the fairly simple to the very complex. For simplicity we may call the ability to deal with such stimuli symbol or abstract intelligence (p. 372).

As to the nature of developmental change in intelligence, Garrett's hypotheses is as follows:

> Abstract or symbol intelligence changes in its organization as age increases from a fairly unified and general ability to a loosely organized group of abilities or factors. If this hypothesis is true, the measurement of intelligence must perforce change in its methods and objectives with increase in age (p. 373).

Holding strictly to Garrett's definition, intelligence does not exist in early infancy. It gradually emerges as the child begins to invest words and symbols with meaning - to let them "represent ideas and relationships." And, of course, this is the reason, according to his view, why baby tests "show little relationship to language tests of school years and hence with symbol intelligence." The evidence cited by Garrett in support of his differentiation hypothesis thus applies only to school age levels when "symbol intelligence" is in evidence, and since intelligence in this sense is non-existent during infancy, the theory has little direct relevance to the problem of infant mental development. However, Garrett suggests that just as differentiation and specialization of bodily movements takes place during infancy, so "this sifting out process is repeated, apparently, in the differentiation of intellectual activities which we encounter later on during the school years" (p. 376). Thus, a process of qualitative change in mental development is clearly implied in Garrett's hypothesis - "intelligence, as I have defined it, changes in its organization with increasing maturity" (p. 373).

At this point it is instructive to compare Garrett's differentiation hypothesis, insofar as it may apply to early mental development, with the theory of Piaget. In Piaget's thinking there is a process which might be called differentiation but conceptually of a different sort. As we have already seen, for Piaget intelligent behavior develops from, or is continuous with, the simple, native reflexes such as the sequence of reflexes involved in sucking. From this point of view, it is not a matter of differentiation of specific patterns from generalized, mass action. There is no "mass" behavior out of which specialized patterns of behavior are "differentiated." The process is rather the "coordination" of specific, ready-made, "sensorimotor schemata" (reflexes) which, in the beginning, being uncoordinated and disorganized, give the impression of generalized mass action. These ready made schemata quickly begin to be exercised and coordinated to form "motor habits" through the "invariant operations" of accommodation and assimilation. The change process is more a matter of "systematization" than of differentiation.

As has already been noted, Garrett sees the development of intelligence (the comprehension and use of symbols) as separate and apart from, and bearing no relationship with, the sensorimotor development of early infancy. For Piaget the reflective intelligence of later ages emerges from, and is developmentally continuous with the overt behavioral development beginning at birth.

Another currently common way of defining the term intelligence is to equate it with the cognitive processes, i.e., perceiving, remembering, understanding. In certain respects this conception of intelligence and its development is similar to that of Garrett. Again, if the term intelligence is confined to this meaning, there would be no intelligence in the very early weeks of infancy. As Ausubel (1958) points out, although the new baby is able to fixate a light patch and follow it with his eyes, and can be conditioned to respond differentially to different sizes and shapes, as well as to different verbal stimuli, there is no real evidence of genuine perception or memory or understanding of the events and situations represented by these stimuli. There is no clear awareness of the properties of external objects or situations. "Hence it is unwarranted to assume that we are dealing here with psychological processes qualitatively similar to perception and cognition in the older individual" (p. 544). In other words, these differential reactions to stimulation constitute psychological experience but of a qualitatively different sort than cognition in which there is of necessity some degree of interpretation in terms of an existing ideational residue from experience. If the meaning of the word intelligence is limited to perception and the other cognitive processes, then, the baby is not born with intelligence. He must acquire it gradually through psychological functioning of a simpler sort. The child gains intelligence only as one acquires the capacity to perceive - not just react differentially to stimuli, but to "know" something of external reality as represented in his sensory experiences.

Such representational awareness of the properties of objects, situations and relationships implies symbolization. Hence the extreme importance of verbal symbols in intelligent functioning in support of this view, vocalizations have been found to be one of the most valid indicators of emerging intelligence (Catalano and McCarthy, 1954). Gesell (1940) stressed the fact that consistent language acceleration is one of the most valid signs of superior intelligence in very young children. R. S. Illingworth (1960) in discussing the reasons for the view that infant tests have no predictive value, pointed out that most infant tests use purely sensorimotor items rather than tests of vocalization, along with items designed to elicit such other signs of cognitive superiority as alertness, responsiveness, and interest in surroundings.

Intelligence as defined in terms of the cognitive processes, or cognitive capacity, since this has not yet emerged in the very young infant, obviously cannot be tested directly at that level. But perhaps, through more careful study of the precognitive behavior of infants, some indicators of potentiality for cognitive development, such as alertness, responsiveness, or perhaps the frequency and volume of random vocalization, could be established.

Still another recently formulated view of the nature of infantile mentality and its development might be labeled "holistic" (Escalona, 1950). In contrast to the point of view described above, intelligence here is regarded as "a fluctuating function, a delicate interaction of many forces and structural conditions" (p. 117). In measurement the objective is to assess psychological functioning as a whole rather than intellectual functioning in the narrower sense. Escalona described her "field-theoretical" point of view as follows:

> A behavior even is determined by the structural conditions prevailing within the organism and in the psychological environment on the one hand, and by forces existing within and without the organism *at this moment* in space and time on the other hand. In other words, the object of psychological inquiry (the test response) is not something located within the infant being tested, but a process of interaction between the infant organism and the immediate environment, i.e., the testing situation. In this theory the testing *situation* becomes the object of analysis quite as much as does the infant's behavior as such (p. 118).

Escalona believes that by basing her assessment upon such a comprehensive analysis of the total "testing situation" the predictive value of the assessment is enhanced Subtle indicators of mental potentiality are presumably sensed and thus constitute in part, the basis for the appraisal. This view also implies developmental continuity without qualitative invariance.

Simon and Bass (1957) agree with Escalona in their statement that a serious and very common error in the evaluation of mentality in infants is to limit attention and consideration to the numerical values, such as the IQ, as a basis for evaluation. They also insist that the real usefulness of testing hinges upon trained clinical judgment in terms of the complete constellation of influencing variables both intra-organic and extra-organic.

Developmental Sequences and the Measurement Problem. Theories of mental development in which qualitative change has been postulated or implied, generally have not been concerned with the lack of correlation between test results in infancy and later school ages. This is particularly true of those conceptions of intelligence which limit the meaning of the term to the reflective of rational processes - the comprehension and use of symbols, cognitive functions, etc. In these instances, the sensorimotor abilities of the infant are regarded as being of quite a different order than the intellectual abilities later to emerge.

However, in every conceptualization of qualitative change which we have reviewed, there is inherent in it the idea of a fixed sequence of "stages" or levels of functioning, each in turn evolving from, or growing out of the previous one. Fundamental to Piaget's theory, for example, is the principle of epigenesis - a continuous sequence of "structures," each qualitatively unique in its organization but developing out of the earlier structure. Gesell's "embryological" view of "behavior patterning" stresses the concept of developmental continuity with qualitative change. Developmental sequences, wherever they are observed - whether it be in the formation of the bones of the hand, or in the growth of mental functioning - are conceived in epigenetic terms. Hence, the fact of a difference in kind does not necessarily mean a lack of relationship. Developmentally, each qualitatively different level of functioning is related to the one preceding it and the one following it. This being true one would expect that valid measures of the various functional levels ("stages") would correlate statistically.

A number of possible explanations for the lack of correlations between infant and school-age tests were suggested above. However, under the assumption of qualitative developmental change, perhaps the most logical hypothesis is that since most of the available tests were not conceived or constructed in terms of observed, qualitatively different levels of functioning in a developmental sequence, they frequently "missed the mark." Constructed as they were on the theory of a general intelligence factor which grows quantitatively, it would be surprising indeed if two tests at different age levels happened to coincide, in terms of content, with the two qualitatively different, but related, functional levels sufficiently closely to measure individual differences and to reveal the actual relationship that might exist between those levels.

Summary - Our Present Theoretical Orientation

In this portion (Part II) of our report we have concerned ourselves with the various theoretical positions, historical and current, in relation to each of three different aspects of the problem of intelligence. These were (1) the question of the determining factors of intelligence, (2) it's nature and structure, and (3) the nature of mental developmental change with time. Related to all of these aspects of the problem is the matter of definition. It is important to make clear in each case just what is meant by "intelligence." The term is used with a variety of meanings, but in what are perhaps its three most common meanings it refers to (1) the genetically determined mental potentiality, (2) the capacity one possesses to acquire new and more adequate modes of behavior and new abilities to function at any particular time, and (3) one's present ability (or abilities) to function, to do, to perform at a particular time.

As to the question of its determining factors, when intelligence is defined as inherent potentiality, it is by definition "fixed" by heredity. The genes determine the possibilities for development. They set the limits beyond which the development of functional capacity cannot go. But even when intelligence is defined as potentiality, its development, nevertheless comes about always and at all levels through the interaction or interchange between organism and environment. Hence, an environment that is not optimum in terms of nutriments, stimulation, or other essentials, may not provide for optimum development. Actually, the full potentiality for the development of mentality at any age level, is rarely reached.

When by intelligence is meant the capacity to acquire increased ability, or more adequate ways of functioning within the limits of inherent potentiality, it obviously is not fixed. Intelligence in this sense, changes with age as the child develops. The rate and extent of this change, likewise, are not fixed or predetermined. The "ceiling" is fixed, but the level of development below that ceiling is not fixed.

When by "intelligence" is meant the ability to perform, its level at any point in time is not only limited by the capacity level so far achieved, but it is also determined in part by past environmental opportunity, occasion, or need to acquire that ability. It is always the ability level, of course, that is observed in a test of intelligence. Any appraisal of mental capacity, or potential therefore, must be arrived at by inference on the basis of what the child was able (or willing) to do. This being true, it could be that two children of the same chronological age, one growing up in a stimulating American home environment, the other coming from an extremely impoverished home in the back hills, actually have equal endowment of intellectual potentiality (intelligence). In terms of capacity for acquisition, for acquiring ability at the time, the two children could also be equal, or, they could be quite different, depending upon relative favorableness of their past environments for optimum mental development. But it would probably be impossible to infer equality, or inequality in ultimate potentiality and present capacity on the basis of intelligence test results alone, because the performance of these children - their ability to cope with the particular tasks of the test (intelligence) might be widely different due to unequal environmental opportunity to learn to perform such tasks.

Theories concerning the nature, or "structure" of intelligence generally have reference to mental abilities, which are conceived as the constituents of intelligence. During the thirty-five year period when most of the currently used tests were developed, the concept of "general intelligence," or "general mental ability" was prevalent. The newer concepts of "primary mental abilities," and of multiple factors of intelligence have come to be dominant during recent years with the development of more effective statistical techniques and electronic computing facilities.

Our present view is that the "structure of intellect" conceived as a dynamic system of distinct, but more or less interrelated abilities, in the normal adult is complex indeed, particularly in view of the possible combinations of "operations," "content," and "products" (Guilford, 1959). But this structure requires time to develop through maturation and experience. The baby's functional repertory is limited to relatively simple operations applied largely, it would seem, to content present to sense (figural content). The baby is usually concerned with the exploration, manipulation and movement of his own body and its parts. He is also sensitive to, and reacts differentially to events and objects outside himself. He plays with his toes, rolls from side to side, reaches for and grasps objects. He turns his head in search of the source of sound. He visually explores a new environment. He differentiates between people. But in this simple sensorimotor activity, mental (central) processes are also involved. He presumably has simple cognitive awareness of the situation. Some memory traces from past experiences must also play a role in determining his behavior. Thus, at this early age, a number of specific mental abilities, not just "general ability," constitute the structure of mentality.

The quality of one's functioning at any age, or stage of development obviously depends upon such specific mental ability factors. One functions effectively in terms of the realities of a situation, only to the extent to which his cognitive abilities make him aware of its various aspects and its complexities, and to the extent that his memory abilities permit him to profit from past experiences, and also to the extent that his abilities to "think" - to try out implicity alternative modes of functioning - facilitate a wise choice of a course of action. In everyday living we "see" ourselves, or judge others as bright, clever, intelligent, or dull, dumb, stupid, always in terms of the apparent adequacy, appropriateness and effectiveness with which a particular life situation or problem is dealt with. In the intelligence test situation, the child is judged as bright, average or retarded, with an assigned mental age or IQ, in terms of the adequacy with which he performs the tasks and solves the problems presented to him.

We prefer, therefore, to apply the term "intelligence," not to innate mental developmental potential, not to the growing capacity to acquire new and more effective modes of behavior within the limits of that potential, not to one's abilities to perform, but rather to the observed, measured or evaluated quality of appropriateness, adequacy, and effectiveness of performance. This quality is, and must always be judged or assessed in terms of what the subject does, or can do.

In devising performance items and constructing scales for appraising this attribute of children's functioning, however, care should be taken, in as far as possible to include items such as to require the exercise on the part of the child, of all the various abilities which modern research has been able to identify as constituting the complex "structure of intellect" in children. Tests and scales that have been constructed on the theory of a single, "general mental ability" may fall far short of the mark.

The question of qualitative, versus quantitative change in mental development also involves the problem of the nature (structure) of intelligence. Again, on the theory of a single general ability factor, development would simply be increase in amount. A child's intelligence would grow by simple accretion. All his new acquisitions, his increasing ability to solve problems, to perform more difficult, abstract and complex tasks, would be based upon, and come about by virtue of this growing, general intelligence factor.

As we have seen, it is difficult to explain the lack of correlation between infant and preschool test scores, and scores obtained during later childhood, in terms of the concept of quantitative change. On the other hand, with the assumption of qualitative change, a feasible explanation is that mentality at different levels in its development is different in its constituent qualities, as well as in amount, and since the tests were designed on the theory of a constant general intelligence factor, rather than in terms of a development sequence of qualitatively different levels of functioning, they fail to register adequately the developmental change taking place. It would seem, therefore, that a promising approach to the construction of mental tests for early childhood might be along the lines established by the work of Piaget (1952), and in the light of findings regarding the structure of mentality as it changes with age (Guilford, 1959). It seems clear that qualitatively different levels of mental functioning do come about in the course of development, and that each level, or "stage," although it emerges from, and is continuous with the preceding one, has its unique and characteristic features in terms of structure. A true developmental sequence of levels of mental functioning, based upon these unique and characteristic features, is needed as a basis for the construction of better infant and preschool-age mental test scales. The final summation of the mental test would, from this point of view, express the levels of mental functioning for a broad band of abilities each of which are possibly developing at different rates depending upon the genetic potentiality and the environmental stimulation for its development. Thus a diagnostic tool would be available for a differential analysis of the various aspects of the child's mental life.

III

INTELLIGENCE TESTING AT THE INFANT AND PRESCHOOL LEVELS

The Development and Evaluation of Tests

> It is said that the most effective tool the artist has is his hands, but that the better able he is to combine skill of hand with the use of delicate tools, the finer and the more expressive his art becomes. The most effective tool the child psychologist has is his knowledge of children, subjective though it may be; but this knowledge functions much more freely and effectively when it is supplemented by tools that enable the psychologist to determine a child's level of development. Any device that renders real assistance in this problem of determining the level of development is of value, but the cruder and more inaccurate tools are of much less value than are refined and accurate ones. Standardized tests of all types fall into the classification of useful tools (Stutsman, 1931, p. 1).

Problems of Test Construction. It is quite obvious that the adequate accumulation and selection of appropriate test items and their accurate scaling depends directly upon the size and representativeness of the sample of children studied at the various ages and levels of development. This presents no particular problem when the concern is with school age children since compulsory school attendance brings together into groups the total population of these children. Thus captive groups can easily be selected to sample adequately the various segments of the population.

The situation is quite different, however, with children under school age. The problem here is largely one of finding the children individually and of getting the parents sufficiently interested to allow their children to be tested. There are, of course, relatively small groups of babies and young children in child-care agencies, day nurseries, etc., that are readily available, but these groups never adequately represent the various segments of the population as a whole. Hence to achieve adequate and representative samples at the various age levels under school age, a great deal of time-consuming, searching activity is necessary.

Another source of grave difficulty in the assessment of mental status in very young children is the extremely limited repertoire of behavior of which they are capable. Since intelligence must always be measured in terms of level and quality of performance, the younger the children to be appraised the more limited the performance possibilities that can be used as indicators of mental status. Furthermore, even though there are wide individual differences in inherent potentiality among infants at birth, differences in immediate behavior possibilities are very small and only become greater gradually as the child grows older. Thus, the problem of finding and devising a sufficient number of test items for test scale construction has been a difficult one from the beginning.

The Binet Tests. For the above reasons there has been until relatively recently in the mental testing movement a marked neglect of the child below school age. The few early efforts, to construct tests at the infant level particularly, did not prove to be very successful. Binet was the earliest investigator to attempt to test children below school age and to establish definite measurement standards. In his first crude scale published in 1905 he included several tests for children under age 3 years. However, there seems to have been little use made of these tests by Binet. In his 1908 revision he included items only for children 3 years of age and older.

In his book published in 1909, however, Binet describes a complete scale, slightly different from the preceding ones, which included seven test items ranging in placement from age 3 months to 2 years. Again these items found little use as a scale. Binet's revision of 1911, and all subsequent revisions by Americans, except the 1922 Revision of Kuhlmann, to be described later, have included no test items for use below age 3 years.

There have been a number of early American revisions and modifications of Binet's scale. Goddard in 1908 published English translations, and in 1911 he published the first American revision of the Binet scale. In this revision no children as young as 3 years were used and only 8 four-year-old children were included. In 1912 Kuhlmann's first American revision appeared. The data used in the organization of this scale came largely from "feeble-minded" subjects. This Kuhlmann himself recognized as an undesirable feature of this revision effort.

Yerkes, Bridges and Hardwick in 1915 produced a point scale using Binet's test items. The scale was designed to make finer distinctions in ability than was possible with the usual form of the scale. Seven-hundred-twenty-nine children were used in the calculation of the norms but only 20 of these were as young as age 4 years.

The scale, therefore, was of little value as a preschool testing instrument.

In 1913 Terman began a thorough revision of Binet's tests. He added some 36 tests from other sources and in 1916 the well known Stanford Revision of the Binet scale was published. This scale also extended down to the 3 year level, but in its standardization only 10 children of age 3, and 51 of age 4 were tested. Thus, as a preschool scale this revision also had very limited value.

One of the earlier evaluative studies of Terman's 1916 revision was made by Mowrer (1933). Her general conclusion was that the scale was not satisfactory at the preschool levels. She found that IQ's obtained at these levels varied widely from test to retest when administered under similar conditions. The intervals she used between test and retest were from 3 to 27 months. For these intervals the IQ changes were found to range between a decrease of 24 points to an increase of 29 points.

Mowrer emphasized the importance of an understanding and appreciation, on the part of the tester, of these wide variations in test results from one time to another in any attempt at interpretation.

Cunningham (1934) made a study of the Stanford-Binet IQ ratings of 37 infants at age 12 months in relation to retest results at 18 and 24 months, and 8 years of age. She found that the incidence of loss of IQ points was somewhat greater in her group than the incidence of gain during the seven-year period. The test administered to this small sample of 12-month-old infants was found to predict ratings at age 8 as well as at age 2 years. The results, however, were regarded as inconclusive due to the small size of the sample.

One of the earliest serious efforts to test intelligence as such in infancy was made by Kuhlmann (1922) with his second revision of the Binet scale. In this revision he extended the series downward standardizing five tests at each of five age levels below age 3 years: 3 months, 12 months, 18 months and 2 years. Mental age based on children's performance on these tests was calculated in the usual way and a table was provided by Kuhlmann for converting mental ages into IQ's. Again, however, an examination of Kuhlmann's standardization data reveals that rather inadequate numbers of children were tested at the preschool levels. The mean mental ages and the IQ's, furthermore are strikingly higher than they should be at these ages. These facts raise doubts as to the validity of any results obtained from that portion of the scale.

Driscoll (1933) investigated the reliability of the Kuhlmann-Binet at the preschool level by retesting a group of 162 "normal" children ranging in age from 12 months to 48 months. Test-retest correlations for intervals of 6 and 12 months respectively were .56 and .76. These reliability co-efficients are obviously too low for individual prediction.

There is no indication in the literature that Kuhlmann's scale was used extensively for measuring mental development during infancy and early childhood.

In 1937 Terman and Merrill published the second Stanford revision of the Binet. It appeared in two equivalent forms, L and M. Much work over a period of nearly ten years went into this revision. It was greatly expanded and completely restandardized. A total of 3184 carefully selected subjects were used in the selection and allocation of items in the scale. An important feature of this revision was the administration of two provisional forms to all subjects with an interval of one day to one week between them. The order in which the two forms were administered was reversed for half of the subjects. Three criteria were used in the final placement of items: (1) percentage of subjects passing at successive age levels, (2) percentages passing the item for each interval of the total score distribution, and (3) the correlation of each item with total score.

This revision was much more adequately standardized at the preschool levels than was the earlier revision. One hundred children were used at each half-year age level from 1½ to 5½ years. At the time of testing every child was within one month of his year, or half-year birthday. Equal numbers of boys and girls were included in each age group.

A study by Bradway (1945), involving 138 children, ages 2 to 5½ years, was concerned with the relative predictive value of the different types of items in the 1937 Stanford-Binet scale. Four types of items, i.e., verbal, non-verbal, memory and number concept, were scored as separate scales and the resulting mental ages were converted into ratios. Ten years later these quotients were correlated with Stanford-Binet IQ's. The results indicated that the verbal and memory scales at the preschool levels were better predictors than the non-verbal ones. Positive, but low correlations were found between original item validity (correlation between item and total score) and the predictive value of the item.

In a later study by Bradway (1958) 111 of the preschool subjects of the standardization group, 1937 revision, were retested as adults after 25 years. A correlation of .59 was obtained between the preschool and the adult IQ's.

During the early 1920's a variety of tests and scales designed for use with infants and young children appeared, most of which were further revisions or modifications of Binet's original scale. Burt's (1921) English revision, designed for children age 3 years and older, was published in 1921. A year later the Herring revision appeared, containing interesting new tests and new adaptations of the old ones. As a preschool scale, however, it was extremely limited since it gave mental ages very little below 5 years. Also in 1922 Trabue and Stockbridge came out with their "List of Typical Performances" for infants in their series of "Mentimeter Tests." This list consisted of tests for children from 3 months to 4 years of age. A year later (1923) a revision of the Yerkes-Bridges point scale appeared. This included a point scale for use with preschool children which had been worked out by Foster (Yerkes and Foster, 1923). The scale, however, extended downward only through age 3 years. In terms of numbers of children tested it is one of the more adequately standardized scales. The first attempt to apply Binet's items to group testing was made by Town.

Her scale published in 1922, consisted of only 13 test items with norms set up in terms of grades rather than chronological age. Although the scale was simple enough for use with 3 and 4 year olds, it was rarely administered to children under age 5 years.

It was also during the early 1920's that the more successful and more adequately standardized infant and preschool tests later to appear were in the early stages of development. One of these was the Linfert and Hierholzer scale for infants which was published in 1928. This scale consisted of two series of tests, one with 29 items for infants from 30 to 132 days of age, and the other of 35 items for children from 153 to 365 days of age. In the standardization of these tests 300 infants were tested, 50 of them repeatedly at different age levels, thus for the first time a fairly well standardized scale designed specifically for the appraisal of mental status during the first year of life was made available.

A few tests not related to the Binet items also appeared during this period, none of which were adequately standardized for children younger than 5 years. Among those were Kohs' Block Design Tests, the Detroit Kindergarten Test, the Rhode Island Intelligence Test, the Pitner-Cunningham Primary Mental Test, and the Cole-Vincent Test.

The Gesell Developmental Schedules. Actually little real effort was concentrated on the problem of assessing mental status in infancy until Gesell began his normative studies in 1919. He and his coworkers began "a systematic survey of the ordinary and normal behavior of infancy and early childhood." The plan was to observe and record infant behavior as accurately as possible for successive age levels during infancy and early childhood and to establish norms for each of these levels. Realizing that "more development is compressed in early than in late stages of the life cycle" their material covering the first year was arranged in "normative schedules" for ages 3 to 4 months, 6 months, 9 months, and one year. The concept of "developmental age" was applied in relation to these normative levels.

Through careful observations of children in their homes and at the clinic, many items of behavior were recorded for the various age levels. The objective was not to construct a scale of general intelligence, or of natural endowment in the strict sense. The selected normative items, however, "always have some significance with respect to developmental status. They are frequently values which show prevailing or predominating tendencies, and as such they serve the purpose for clinical orientation, for orderly observation, and interpretation of both capacity and personality" (Gesell, 1925, p. 44).

From this material the Gesell Developmental Schedules were prepared. These schedules were simply categories of items arbitrarily arranged to allow a rough appraisal of the child's level of development in each of four major areas: motor behavior, adaptive behavior, language behavior, and personal-social behavior. The procedure in the use of these schedules is purely observational. In general, they constitute a standardized procedure for observing behavior and evaluating the course of behavioral development.

The Gesell schedules have been one of the main sources of material in the construction of subsequent tests for young children.

As was true with other tests of infant ability, evaluative studies of Gesell's schedules generally have demonstrated practically no correlation with later tested intelligence. The few factor analysis studies which have been done suggest a changing nature of "mental ability" from time to time during infancy. It was not clear in these studies however, to what extent these changes in factor content and meaning reflected actual qualitative change in intelligence or to what extent the apparent change might have been a function of time item content of the schedules.

Richards and Nelson (1938) in a factor analysis study of results from the six months Gesell Schedule found three factors which they called testability or halo effect, alertness and motor ability. In a later report (1939) of studies at the 6, 12 and 18 month levels, they described the three factors at the six months levels as a general factor, alertness and motor ability. Beyond the six months level, however, only two factors were clearly in evidence. These were a motor factor involving both gross motor activities and finer manipulations, and "alertness." It was found that the items signifying alertness correlated best with later mental status. Their observation was that with advancement in age this alertness factor became a more and more important element in mental tests.

In 1960 Knobloch and Pasamanick reported a study of consistency and predictive value of the 40 week Gesell Developmental Schedule. As a result of their experience with the schedule, they regarded it as a "neurologic examination" stemming from the fact of the "integral interrelationship between neurologic integrity and maturational development." The study was based on the evaluation of 500 premature infants and 492 full term "controls."

They reported correlations of .98 between the results of different examiners with regard to general developmental quotient, neurologic status and intellectual potential. They also reported a correlation of .87 between the 40 week Gesell and the Stanford-Binet at age 3. Their general conclusion was that the Gesell Developmental examination is "a reliable and valid tool for identifying infants with deviations in neurologic or intellectual development."

The Merrill-Palmer Scale of Mental Tests. During this period of the 1920's most attention was given to the later preschool years. Stutsman (1931) since 1922 had been at work on what was to be the Merrill-Palmer Scale of Mental Tests for Preschool Children. This scale was published in 1931 and has been through the years one of the most widely used preschool scales.

To serve as a guide in the selection and construction of test material Stutsman formulated a set of 12 criteria. These were as follows:

1. The test material should have inherent interest for the child.

2. A large variety of activites and abilities should be tested.

3. There should be variation in difficulty in test items.

4. Performance on the test should be influenced as little as possible by the child's training and environment.

5. Simplicity of material is an asset.

6. The test should be easy to administer.

7. The methods of scoring should be objective, eliminating subjective judgment as far as possible.

8. The test items should markedly differentiate between the age groups of children.

9. The age range of each group should be small enough to prevent a high positive correlation between the score and the age of the child within the group.

10. The selection of cases should be as near a sampling of the whole community as possible.

11. The standardization group should be large enough to render the conclusions valid statistically.

12. The final scale must be easy to score.

In terms of these criteria 38 tests yielding a possible total of 93 scoreable test elements were selected for the final scale. An "age at par" (age at which 50 per cent of the normative sample passed it) was computed for each test element. The items were then arranged in the ascending order of difficulty according to their ages-at-par. They were also grouped into 6 month age levels. The point-scale system of scoring was adopted. From these raw scores certain derived scores, including the mental age, can be computed.

A total of 631 children ranging in age from 17 months and 16 days to 77 months and 15 days were in the standardization.

As evidence of the validity of the scale, the following summary points were made (Stutsman, 1931, p. 116):

1. The test elements were selected originally because they differentiated children who, as judged by the nursery school staff, showed differences in mental ability.

2. The total score showed a correlation of +.92 with chronological age.

3. The overlapping of the distributions of total scores between age groups was relatively small.

4. The correlations with the Stanford-Binet scale for three separate groups were .79, .79 and .78.

No information on reliability is given in the manual. However, test-retest correlations later reported ranged between .77 and .96.

A study by Driscoll (1933), however, reported somewhat lower reliability coefficients (.47 - .60). Although Stutsman had shown that the IQ could not be appropriately used with the Merrill-Palmer Scale, these test-retest correlations were based upon the IQ's of 73 boys and 85 girls through an age range of 12 to 48 months. The interval between tests, however, was 6 to 12 months. Even though the IQ is a ratio, and is thus presumed to nullify the passage-of-time factor, in view of the changing patterns and rates of individual development, and the widely varying effects of environmental stimulation, especially during this period of rapid growth, such a test-retest correlation with an interval of 6 to 12 months can hardly be regarded as an appropriate index of the reliability of the testing instrument. DeForest (1941), correlating standard deviation scores instead of IQ's, obtained a test-retest correlation of .60. The interval between tests in this instance was 4 to 24 months.

With the assumption of "fixed intelligence" and genetically predetermined mental development it was felt that a test which was adequately standardized at the various age levels should show relative constancy of IQ. In some studies of the Merrill-Palmer the test-retest procedure was used with the specific purpose of evaluating its standardization. Kawin (1934) was concerned with this question in her study of 169 children. They were first tested with the Merrill-Palmer at ages 23 to 57 months (mean age, 34.1), then retested at 28 to 64 months (mean, 43.6 months). The test-retest correlation was .59. Kawin's conclusion was that the Merrill-Palmer scale was not satisfactory for purposes of prediction because of its inadequate standardization. Quite a different opinion was rendered by Teagarden (1940) when she wrote that the Merrill-Palmer scale must be considered one of the best available scales because of its careful standardization. She pointed out that because of children's interest in the activities of the scale, the skilled examiner has no difficulty getting cooperative responses from the children. It is especially valuable, she felt, in testing older children who, for undetermined reasons, do not talk.

Mowrer (1934) made an evaluative study of the Merrill-Palmer scale with a somewhat different approach. She selected 18 test items from the scale for special study - from the point of view of their power to discriminate varying degrees of mental ability at the various preschool levels. She used the point scale method of scoring recommended by Stutsman and compared each child's score on each test item with the percentile scores for his age level. Because of the small number of cases in her study (83 preschool children), Mowrer made no direct comparisons of her results with Stutsman's percentiles. Her general conclusion, however, was that when the results obtained from the tests she studied were combined with data from other tests, "they furnish a valuable source of information in the study of individual children of the preschool ages" (p. 96).

As we have seen, most of the evaluative studies of infant and preschool tests made during the 1930's and 1940's were concerned with the question of predictive validity. The most frequent criterion used was later Stanford-Binet IQ. Driscoll in her 1933 study reported a correlation of .61 between the Merrill-Palmer administered between 2 and 3 years of age and the 1917 revision of Stanford-Binet at "5 years of age or over."

The Stanford-Binet criterion was frequently used in evaluations of other aspects of validity than predictive power. Black (1939), for example, using small groups of nursery school children from families of high social status, found different degrees of correlation, between the Merrill-Palmer and the 1937 Revision of Stanford-Binet administered at the same time.

The children varied in age and had had different amounts of nursery school experience. Thirty-one "new" children, ages 24-54 months scored somewhat higher on the average on the Stanford-Binet than on the Merrill-Palmer (Mean IQ's 121.7 and 115.6). The correlation between the two sets of IQ's was only .36. For children 6 months older and with several months of nursery school experience, the mean Binet IQ was 131.2, mean Merrill-Palmer IQ was 129.8 and the correlation between the two was .64. Black concluded on the basis of his studies that the Merrill-Palmer is "less stable" than the 1937 Stanford-Binet.

DeForest (1931) obtained a correlation between these two scales of .55 when all her age groups were combined. With smaller age ranges, however, she discovered a relationship between age and the magnitude of the correlation. For her three age groups, 24 to 39 months, 30 to 45 months and 36 to 51 months, the respective correlations with Binet were .71, .65 and .65. For her 51 to 63 month group the correlation was only .22.

By use of the Wherry method of analysis, Hurst (1960) factor analyzed the Merrill-Palmer scale. Tetrachoric item-intercorrelations were computed from the test scores of a group of three-year-old nursery school children. The factor analysis resulted in six factors which Hurst called "willingness to cooperate," "finding relations," "fine motor coordination," "persistence," "perceptual speed" and "space." He stated that only three of the six factors seemed to be "intellectual" in nature and with characteristics common to those found in tests of later mental ability. There were the factors he called "finding relations," "perceptual speed" and "space." The remaining "extraneous" factors, Hurst felt, were largely the basis for the scale's low predictive validity.

The Minnesota Preschool Scale. During the late 1930's the Minnesota Preschool Scale was in process of being developed by Florence Goodenough and her co-workers (Goodenough, et. al., 1932-1940). This scale is available in two equivalent forms each consisting of 26 items. Many of its test items were adopted from the Kuhlmann-Binet scale, some from other available scales, still others were new and original.

Nine-hundred carefully selected children constituted the standardization sample. These came from nursery schools and public and private schools. There were in the sample 100 children for each half-year age level from $1\frac{1}{2}$ to 6 years.

The scale is made up of both verbal and non-verbal tests with relatively few motor items.

Norms are provided for each of the half-year levels, for the verbal and non-verbal items separately and for both combined. These can be expressed as IQ equivalents and also in terms of equal unit scale scores.

Reliability coefficients for the verbal scale ranged from .68 to .94 and for the non-verbal scale, .67 to .92. For the total scale the range was from .80 to .94.

Validity coefficients from follow-up studies with the Stanford-Binet Scale as the criterion, varied considerably depending upon the ages when the tests were administered. For example, the median correlation between the Minnesota Preschool IQ's obtained at 4 years and older, and Binet IQ's obtained when the children were between ages 4½ and 13½ years was .68. In cases where the Minnesota scale was administered before age 3 the correlation was only .21.

Maurer (1946) made an adult follow-up study in which 226 subjects who had been tested with the Minnesota Preschool before the age of 6 years were located and tested with the Army Alpha. They ranged in age at the time of the adult test between 16½ and 22 years. Correlations between the Minnesota test, Forms A and B, and the adult Alpha for a subsample of 46 subjects were .31 and .33 respectively.

In general, among the strong points of the Minnesota are its careful standardization, its two parallel forms, with high inter-form reliability. On the negative side the tests are not particularly appealing, especially to children under three years of age.

The Goodenough Draw-a-Man Test. A number of performance and non-language tests have been devised, some of which extend well into the preschool age range. Perhaps the best known of these is the Goodenough Draw-a-man test (Goodenough, 1926). The child is simply instructed to "make a picture of a man; make the very best picture you can." Emphasis is placed upon accuracy of observation and the development of conceptual thinking rather than upon artistic skill. The test is scored in terms of such items as the parts of the body included, clothing details, and perspective. In 1961 a revision and extension of the test was published by Harris. In this revised form the child is asked also to draw a picture of a woman and of himself. The Self scale was included more as projective test of personality. A sample of 300 children was used in establishing norms for the Man and the Woman scales at each age level. Since these scales project downward only to age 5, they have limited usefulness as preschool tests.

In a study by McCarthy (1944) two main sources of unreliability of the Goodenough Drawing test were described. One of these is the change which frequently takes place in the quality of the child's drawing as measured by the scale on two different occasions. The other source of unreliability is that due to subjectivity of judgment on the part of the scorer of the test.

The California First Year Mental Scale. By 1933 Bayley had accumulated a considerable number of test items from various sources, particularly from Gesell, Stutsman, and Goodenough. With many of her own added, and by adopting many of the available ones for lower age levels, her published "California First Year Mental Scale" contained 115 test items covering a mental age range from birth to 18 months. An attempt was made "to compensate for the reduced reliability of tests of very young children by including a larger number of items than are usually regarded as necessary." (p. 5) Effort was also made to provide continuity of content by using appropriate tests of the same functions at successive age levels, a condition so lacking in most intelligence scales.

The numbers of test records used at the various age levels in the standardization of the scale were not large, ranging between 52 and 61. However, a feature of this work rarely found in standardization procedures is the longitudinal nature of the data. The same infants (52 to 61) were tested monthly, yielding some 870 separate test records. Thurstone's absolute scaling procedures, with standard scores for each age level were used. Split-half reliabilities ranged between .51 at age 2 months, to .95 at age 6 months with a median value of .86.

In Bayley's, (1933, 1955) longitudinal study with this scale, however, she obtained correlations between tests administered under one year of age, and retests at 18 months, of practically zero. With later retests the correlations became negative and as high as -.21.

The published scale was provided with full directions for administering and scoring, and with age norms and reliability coefficients. This scale was quite widely used, but, along with other infant scales, became generally discredited because of its failure to predict later Stanford-Binet IQ.

The Cattell Infant Intelligence Scale. This scale was regarded by many psychologists as the most satisfactory device for testing very young children (Cattell, 1940). The scale was standardized from a longitudinal study of 274 children. Because the entire group was not available for all of the retests, the actual numbers upon which the norms are based vary considerably. The children were tested at ages 3, 6, 9, 12, 18, 24 and 30 months.

The items for the final scale were selected in terms of several criteria. Among these were ease of administering and scoring, objectivity, lack of cumbersome equipment for administering, attention-getting and holding value, and increase in the percentage of passing from age level. Tests requiring a time limit were deemed undesirable for use with very young children. There are no time-limit tests in the Cattell scale.

The test materials are very similar to those used in the Gesell Developmental Schedules and the lower levels of the Binet scale.

Validity indices based upon correlations between Cattell IQ's at different age levels and Stanford-Binet, Form L, IQ's at age 3 years range between .10 for the 3 months age level, and .83 for the 30 month level. Thus, there is a lack of predictive value in the scale, in terms of the 3-year Stanford-Binet IQ for ages under one year.

Escalona (1950) made the first published evaluative study of the Cattell scale. She found that conditions under which infants were tested had much to do with the predictive value of the results. Under what she described as "optimal conditions" the variation between test and retest tended to be less than when conditions were "non-optimal." Escalona's general conclusion was that for best results the "gestalt" of the total testing situation must be taken into consideration rather than the mere test finding itself. Gallagher (1953) arrived at essentially the same conclusions. She felt that it was not safe to place too much reliance on a "useful but not overly reliable instrument."

Cavanaugh (1957) concerned herself among other matters, with the predictive validity of the Cattell scale. Her findings were in accord with those of earlier investigators, that intelligence scores of children under two years of age are of little value in predicting subsequent IQ scores. She did find, however, that the scale discriminated significantly between premature, mature, and post-mature infants at ages 6, 12, 18 and 36 months.

The Northwestern Infant Intelligence Test. Gilliland (1948) reported the development of a scale for the measurement of infant mentality which had been in process of construction and standardization during the preceding six years (Shotwell and Gilliland, 1943). In the face of much published evidence that in general the measurement of infant intelligence has no predictive value, Gilliland and his co-workers began the task of assembling a set of "predictive items which measure adaptation to physical and social environment." Items were selected from the available tests, mainly the Gesell Schedules, and other items were adapted or devised, resulting finally in a scale of 40 items, covering age levels of 4, 8 and 12 weeks. The criteria in terms of which the items were finally chosen were: (1) there must be definite directions for giving and scoring each item, (2) they must require only relatively simple equipment, (3) they must "measure adaptation to the environment rather than physical growth or maturation," and (4) they must show definite age-grade progress.

The final revision was based on the test records of 276 babies. The overall reliability of the scale was .84 and there was "evidence of high validity."

Early Performance Scales. The idea of measuring mentality through motor performance goes back to the early days of the testing movement. In a sense, every "mental" test is a "performance" test since there is no way of appraising mental ability except in terms of what the subject can do. In the "performance test," however, the effort is made to eliminate, in as far as it is possible, the use of language on the part of either the examiner or the subject.

The first major attempt to standardize a series of such tests in the form of a scale resulted in the Pintner-Patterson Performance Scale published in 1917. All of the tests in this scale were presented without language and no language was required in the responses. It consisted of 15 tests, many of which have been included in more recently standardized scales. A later shortened form contained 10 of the original 15 tests. A number of methods of scoring and calculating mental age were suggested by the authors. The reliability of the Pintner-Patterson scale proved to be considerably lower than that of most verbal scales.

In 1930 Arthur released for use the Arthur Performance Scale. This scale contained eight of the Pintner-Patterson tests, together with Porteus (1924) mazes and the Kohs (1923) Block Design test. All of these tests were restandardized on a new sample of 1100 children, ages 5 to 15. A "Revised Form II" of the Arthur was published in 1947. Because of its limited use at the preschool age levels the Arthur scale has only historical significance to our present topic.

The Leiter International Performance Scale. In recent years a number of attempts have been made to devise non-language mental tests such that the responses to them are not influenced by experience or learning peculiar to any culture. Certain of these so called "culture free" or "culture fair" tests have been standardized for children as young as two years. The Leiter International Performance Scale (Leiter, 1950, 1959), is an outstanding example of this type of instrument. The scale consists of a series of tests which were developed over a period of years of use with different ethnic groups in Hawaii and Africa. A revision of the scale was issued in 1948 largely based on testing of American children. A distinctive feature of the scale is the elimination of instructions, by pantomine as well as spoken, and the comprehension by the child of the task to be performed is treated as part of the test.

The Leiter scale was designed to cover much the same functions as those tested by the usual verbal scale. Among the tasks set for the child are the various matching tasks - colors, forms, pictures, picture completion, copying designs, analogies, series completion, recognition of similarities and footprints, immediate memory, classification of animals. The 1948 revision extends down to age two years.

The scale is administered individually and there are no time limits. It provides both mental age and IQ scores. Reported split-half reliabilities range between .91 and .94. Correlations with Stanford-Binet scores are reported to range between .64 and .81.

Tate (1952), in her comparative study of the Leiter, the Arthur, and the Stanford-Binet, Form L administered to 108 five-year-old children, concluded that when comparable scores, such as the standard score, were used the three scales were, "at least in part, testing the same thing" (p. 501). The correlation between Leiter and Stanford-Binet was .81. According to Tate's test of freedom from cultural effects of the three scales, however, the Leiter was found to be no more culture free than either the Arthur or the Binet.

In 1952 Arthur published an "adaptation" of the Leiter scale. This revision was standardized as a point scale and is suitable for use with older preschool children.

The Full Range Picture Vocabulary Test. In all mental tests for young children, with exception of the non-language tests, vocabulary items have been used rather extensively. Tests of vocabulary generally have been found to be valid indicators of mental developmental level. A scale published in 1948 by Ammons and Ammons is unusual in that it is a vocabulary test which is essentially non-verbal in the sense that the "testee need not say a word." He is simply asked to point to one of four pictures on a card which best illustrates the meaning of a given word spoken by the examiner. Sixteen such cards, each with a list of words graded in difficulty constitute the scale. It requires only 15 minutes to administer.

The standardization of the scale as a preschool instrument (Ammons and Holmes, 1949) was based on a "representative preschool-age population" - a sample of 120 American born white children ranging in age between 2 and 5 years. This sample included the same proportions from the various occupational groups as was found generally in fathers of preschool children of the United States in the 1940 census.

The scale was published in two forms with a reliability of .93 for the preschool age range. Its correlation with the 1937 Stanford-Binet was found to be .85 for Form A and .83 for Form B.

The Griffiths Mental Development Scale. In 1954 Griffiths reported a new scale designed to measure infant mentality. The total scale consists of five sub scales presenting a "balance" of five different areas of behavioral development, namely, the locomotor, the personal-social, hearing and speech, hand and eye development, and performance. Each of these subscales contains fifty-two "finely graded items arranged in order of difficulty," making a total of 260 items in the whole scale. It is designed to cover the first two years with three items for each week of life for the first year and two items for each week of the second year. The purpose in the construction of the scale was,

> ....to develop a measuring instrument that shall be at once reasonably exact, well balanced and comprehensive, so that those activities that later go to make an efficient person, or as many as possible of those skills that are measured in assessing intelligence in older subjects, shall be represented, as it were in embryo, in a valid test of the mentality of a baby (pp. 31-32).

Griffiths conceived of intelligence as general ability, or "g" and she set out to measure as many as possible of its various manifestations. She borrowed many items from other scales, especially from the Gesell schedules, but she devised many more of her own. Two test-retest reliability coefficients were reported for the scale. These were .87 based on 60 cases, and .92 based on 52 of the 60 cases after the 8 cases whose histories showed "disturbing or unusual conditions" were eliminated. No claims were made as to the predictive validity of the scale.

The A. Brenner Gestalt Test. Brenner (1959) has recently devised a simple, and easily administered test for appraising children's "readiness" for entrance into the first grade of elementary school. The test consists of five tasks. "In their basic patterns these tasks are built on the principle of Gestalt. The child must either recognize and reproduce or produce a particular Gestalt" (p. 3). They are (1) copying ten dots of equal size arranged in a specific order, (2) copying the sentence, "Fred is here," the instructions varying with the age of the children, (3) drawing a man, (4) number producing activity, using one-centimeter cubes, and (5) number recognition activity. "The major concept is that perceptual-conceptual development is a main agent in personality development, learning and readiness for school" (pp. 23-24). The total score on the Brenner Gestalt test was found to be correlated to the extent of .81 with teachers' ratings of the children's levels of functioning.

A New Revision of the California First Year Scale. "An

Infant Scale of Mental Development," covering age levels from birth to 15 months, has recently been accomplished by Bayley and her co-workers. The new series of tests was developed by using the original California norms as a frame of reference but with alterations, deletions and additions. The new items have been "borrowed or adopted" mainly from the scales of Griffiths and Cattell.

Approximately 1000 testings of children of one month through 15 months of age furnished the new standardization data. These were obtained from the examiners of the National Institute of Neurological Diseases and Blindness Collaborative Research Project.

In Bayley's complete examination of a child, the following instruments are used: (1) the Scale of Mental Development, (2) the Scale of Motor Development, (3) the Baby's Day Interview, and (4) the Infant Behavior Profile. For the standardization tests, an additional part called "Classification Information" is also to be filled out by inquiry of the mother while she is with the baby. The Baby's Day interview and the Infant Behavior Profile, furnish information supplementary to the test scales and are to be considered integral parts of the total developmental scale.

The point scoring system is used. The point scores are then converted into standard scores by use of a table of means and standard deviations. These norms are provided for each half-month level from one month to 15 months. Mental ages, of course, may also be obtained from this table. However, the relatively small standard deviations at these early ages (narrow ranges in mental age for given chronological ages) make the use of the IQ inadvisable.

This new set of scales, and the assessment procedures associated with them, undoubtedly constitute the most adequate mental testing program at the infant level yet to appear. As of this writing, they have not been made available for general use.

Published Studies which Involved the Testing of Infants and Preschool Children

Studies Concerned with the Nature of Early Intelligence and the Factors in its Development. Earlier in this report, considerable attention was given to the general finding that infant test results are of no practical value in predicting later levels of mental functioning as measured by tests of intelligence. Reference was made to the many studies in which the lack of predictive validity was demonstrated. During more recent years a few studies of a more analytical nature have appeared in which attention has been directed toward the identification of the factors responsible for this state of affairs. Escalona (1948), for example, unable to accept the dictum that infant tests have no predictive validity, concluded on the basis of her study that test results when adequately interpreted in the light of concomitant clinical observations do have predictive validity. She found that certain aspects of mental functioning can be predicted more accurately than others. In later articles (1950, 1954) Escalona elaborated on her "field-theoretical" procedures for the administration and evaluation of infant tests. She emphasized again that the predictive accuracy of such tests can be improved by giving special attention to the qualitative aspects of mental functioning in relation to the results of clinical observation in the context of the total testing situation.

Illingworth (1960, 1961) also questioned the generally accepted conclusion that infant tests have no predictive value. He enumerated several reasons for this negative view, one of them being that in the selection of cases for testing for predictive validity, mentally retarded children are generally excluded in spite of the fact that the detection of mental retardation and neurological conditions has always been one of the main functions of infant mental testing. Illingworth's evidence shows that these conditions can be readily detected in early infancy by the use of infant tests. Escalona (1950) also points out that infant testing has succeeded in detecting the mentally deviant at very early ages.

In our earlier discussion of the nature of developmental change in mentality it was suggested that the lack of correlation between infant test ratings and those obtained at later school age levels might possibly result from the fact that infant intelligence is different in kind, as well as in amount, from intelligence at older age levels. Bayley (1954) in her study of parent-child similarities furnishes some evidence of such a qualitative difference. She obtained a zero correlation between intelligence ratings of parents and those of their children during the first 15 months. After that age, however, the correlations rose rapidly to +.50 at age two years. Assuming that all of the tests involved in the study were really measuring "intelligence," the suggestion would be that intelligence in early infancy is not of the same quality as school age and later intelligence.

Another area of interest to researchers which involves the use of infant tests, has continued to be the question of environmental factors in relation to stability or change in IQ ratings. One of the earliest studies in this area was that of Goodenough (1927) in which she found that test scores of preschool children correlated positively with the educational level of parents. A little later Van Alystine (1929) found a similar relationship between intelligence measures at age three, and parents' occupational ratings. On the other hand, Furfey and his associates (1928, 1932) found that the mental "ability" of infants as measured by the Linfert-Hierholzer scale was not significantly related to the socio-economic status of the parents.

Bayley and Jones (1937) provided "a link between these discrepant findings" in a longitudinal study of this relationship in two larger groups of 405 and 59 children respectively. The environmental factors studies were parents' education, family income, father's occupation, and "social-cultural-economic factors represented in the Berkeley Social Rating Scale." All of these factors showed zero or slightly negative correlations with mental test scores up to 18 months of age. After this age the correlations increased to significant levels. The factor that gave the highest correlations was mothers' education (.50 at children's age 5 years). In contrast to these results, Levinson (1959) in his study of bilingualism found no correlation in his groups of 4 to 7 year old children between socio-economic background and measured intelligence.

Peters and McElwee (1944), working with nursery school children, were convinced that a special teaching regime which they devised, based upon an analysis of intelligence into its components, was of practical value in improving the "functioning intelligence" of the children. The subjects of the study were six children, 2 years 9 months to 4 years 2 months of age, with tested IQ's below 100, and who had been rated by their teachers as low in initiative and in "quickness to learn." They all came from environmentally deprived homes. At the end of a period of 8 months, during which the children had been given the special program, there was a mean gain in IQ of 7 points. These findings, being based on only six children, were obviously not statistically significant.

Another study of special environmental circumstances in relation to early tested mental facility was that of Klatskin (1952). He tested a group of 316 infants who, with their mothers had been housed during their first year of life in a rooming-in unit under an unusually lenient and permissive regime and excellent pediatric care. At 12 and 13 months of age these infants on the average scored significantly higher on the Cattell Infant Scale than did Cattell's standardization group. Although Klatskin was very cautious in the interpretation of his results, the suggestion was that the exceptionally "favorable" environment in which the babies developed was at least one of the factors operative in the outcome.

Bradway (1945) used the Stanford-Binet scale in a study of IQ changes in relation to background factors. Her subjects were a group of children whose IQ's had increased on the average by 9 points between preschool and junior high school levels, and a group from the same population whose mean IQ had decreased by 11 points during the same period. The environmental backgrounds of these two groups were studied to determine whether there were any factors which differentiated them. The adequacy of mothers' vocabulary, parents' intelligence and ancestral intelligence were found to be significantly differentiating factors.

Two rather impressive studies with a biological orientation and emphasis, in which the Gesell Developmental Schedules were used, were reported by Knobloch and Pasamanick (1959, 1960). In the first of these, 992 infants, about half of which were prematures, were examined and tested at a mean age of 40 weeks. The prematures and the full-term control infants were matched on economic status and educational level of parents, and race. They found a definite relationship between birth weight and developmental quotient for both white and non-white babies. This finding is in disagreement with an earlier report by Benton (1938) who, in his study of 78 prematurely born children, found no evidence of adverse effects of prematurity on tested intelligence at preschool ages. Knobloch and Pasamanick also found no significant difference in distribution of DQ's between whites and non-whites. The main point of emphasis in both studies was the basic importance of "neurologic status" - the "substrate of neural organization" which they felt not only operates to mold behavior in infancy but is the foundation of all learning. The Gesell Developmental Examination was regarded as an extremely valuable device for diagnosing the "integrity of the nervous system" as well as the general developmental level of the infant.

The relationship between the early manifestations of intelligence, and speech sounds and language development has also been an area of interest studied by the use of mental tests. Spiker and Irwin (1949) found a small but reliable relationship between certain aspects of infant speech sound development (phoneme type and vowel sound frequency) and the abilities measured by the Kuhlmann Test of Mental Development. Anastasi and DeJesus (1953) made a study of the language development and non-verbal IQ based upon the Goodenough Draw-a-Man test in Puerto Rican Spanish speaking children, and white and negro children of New York City. These children were all within 6 months of their fifth birthday. Although their parents were below those of the other two racial groups in educational and occupational level, the Puerto Rican children did not differ significantly from the others in non-verbal IQ, and they excelled both white and negro children in mean sentence length and in maturity of sentence structure. The greater extent of adult contact in the homes of the Puerto Rican children was suggested as one possible factor to account for their superiority in early linguistic development. Such evidence, coming from two quite different preschool testing approaches, demonstrates the differentiating roles played by the various aspects of the social environment in the development of mentality. It also lends support to the view that verbal and non-verbal "intelligence" are "factors" that do indeed vary somewhat independently of each other at the preschool level.

Use of Intelligence Tests in Early Childhood for Research Concerning Other Psychological Variables. Infant and preschool tests have been used for the investigation of theoretical problems other than those concerned with the nature or development of intelligence *per se*. Koch and Meyer (1959), for example, studied the relation between intelligence and the formation of "learning sets" (rates at which simple discrimination learning is transferred from one problem to another) in preschool children. The children ranged in age from 2½ to 5½ years with an IQ range of 96 to 146, mean 115. Koch and Meyer found a wide variation in amount of time required to reach their criterion of learning. The correlation between time (number of days) and mental age scores was -.50. Thus, according to these findings, learning set formation varies with Binet intelligence at the preschool level.

Moriarty (1960) used preschool intelligence tests, not only to assess the child's general level of mental functioning as indicated by mental age or IQ, but especially to study functioning intelligence "as it is implemented, modified or controlled" by affective factors, and particularly by the situation with which the child is coping as he responds to the test. In general she found that patterns of behavior observed in the structured test situation closely resemble those observed in other situations.

Especially interesting was her finding that ratings of non-cognitive aspects of the child's behavior, such as the feeling of pride and confidence in oneself, or being able to accept one's limitations, were more closely related to general effectiveness in coping than were the IQ scores. If by "intelligence" is meant effective coping with life situations, then according to Moriarty's findings a single numerical score (MA or IQ) is not necessarily the best index of intelligence.

The Stanford-Binet was used in an "exploratory study" by Schafer and Leitch (1948) to appraise its usefulness in detecting maladjustment tendencies in children 2 to 5 years of age. The mental test was used in conjunction with psychiatric interviews, the Rorschach and the first 10 cards of the Murray Thematic Apperception Test. It was possible for the psychologist, on the basis of test results to distinguish three groups of children rated by the psychiatrist as "adequately adjusted," "moderately adjusted" and "severely maladjusted." The degree of uneven functioning on the items of the Stanford-Binet Test, as measured in terms of the scatter of successes above the mental age and failures below the mental age, was found to be a useful indicator of adjustment level. Other qualitative analyses of the test results were found not to be fruitful at the preschool age levels.

From their longitudinal data on 39 children, Despert and Pierce (1946) made a study of the relation between emotional adjustment and intellectual functioning as measured by the Terman-Merrill revision of Stanford-Binet, during the children's nursery school attendance.

In 22 of the 39 children, significant changes ranging from 10 to 31 IQ points up or down were recorded. It was with this group showing maximum IQ change that a significant relationship with emotional adjustment, as measured in terms of the total mass of data, was found. Again in this study, the qualitative analysis of the test items did not reveal significant patterns nor reliable criteria for emotional maladjustment, although scatter up and down from mental age in item response did have significance. The analysis of their total records, including anamnestic data, indicated that the children's "total emotional adjustment" influenced their psychometric scores.

One of the persisting perennial problems that beset school administrators is that of initiating and properly orienting the hordes of six-year-olds who present themselves each year for entrance into the regular elementary schools. One obvious need is some sort of fore-knowledge as to each child's degree of readiness for the public school experience. There are, of course, a number of "reading readiness tests" in wide use, but relatively little is being done in the way of a comprehensive assessment of total readiness for school experience. Any program designed to make such an assessment would involve the use of some form of preschool mental testing. Relatively few published reports of this use of tests are to be found in the literature.

Grassi (1955) published a report of a program for assessing school readiness prior to the child's entrance into the elementary school. The program was based on the idea that a child must have reached a minimum stage of maturity - mentally, socially, physically and emotionally if he is to meet school expectations in the first grade. It involved in each case a visit of the child and his mother to his prospective school prior to the opening of school, where, in addition to being greeted and made to feel welcome, he was placed in a group of his peers and a "mental maturity" test was administered to him. The results of the study indicated that for children who had attained a mental age of 6 years or more on a psychological test, the accuracy of prediction of their success in school was 97 percent. In those cases where the test indicated a mental age below 6 years, the correct prediction was only 93 percent. The total overall prediction accuracy in this study was approximately 95 percent.

Brenner (1956, 1959) in his investigation of readiness for school, and his search for criteria of readiness, has used a number of preschool tests including the Goodenough Draw-a-Man, the Sangren Information Test for Young Children, and the Pintner-Cunningham Primary Mental Ability test. As a validating criterion he used teacher's ratings of the children's level of functioning in the second half of the first grade, combined with the teacher's rankings in terms of the various tests given earlier. The rank-order correlations of the mental tests with this criterion were as follows: Goodenough .69, Sangren .76, Pintner-Cunningham .68. These correlations were significant at the 1 percent level of confidence, and high enough to indicate that mental ability as measured by mental tests, as has long been assumed, is perhaps in general the most basic single factor in children's readiness for school.

The majority of published studies of the relation between emotional or personality disturbances are concerned with children beyond the preschool age level. In this research area, as well as in the use of infant and preschool tests for practical purposes, such as mental assessments for admission or placement in schools or other institutions, relatively few published reports are to be found.

IV

QUESTIONNAIRE SURVEY OF THE CURRENT SITUATION WITH RESPECT TO INFANT AND PRESCHOOL MENTAL TESTING

In our evaluation of the overall situation with respect to the assessment of intelligence at the infant and very early childhood levels, for basic information we turned to those on the "firing line," those individuals who, with one purpose or another, are actually using the available infant and preschool-age tests. We were interested in the relative frequency with which each of the various tests are presently being used, and for what specific purposes. We were particularly interested in the judgments of the users of the tests as to their value, effectiveness, and limitations, and in what ways the users felt the tests could be improved.

A questionnaire designed to obtain such information was drawn up. After a trial as an interview instrument, the questionnaire was revised and somewhat enlarged. A copy of this revised three-page questionnaire may be found in the appendix.

Result of the Survey

Copies of the questionnaire were mailed out to some 750 carefully selected names. These included research centers, treatment agencies, and individual clinicians, practicing psychologists, and researchers. Three-hundred-thirty, or 44 percent, of the questionnaires were filled in and returned. Of these, 113 indicated that infants and preschool children were not included in their clientele. The results herein reported, therefore, are based on the 217 returns reporting service to, or work with, children under six years of age.

Services to Young Children and Their Parents. Our first querry was concerned with the kinds of psychological, and other services currently being made available to children and their parents by those who do testing, as represented by our respondents. (See Table 1.)

Ninety-three percent of the group specifically reported mental testing of infants and (or) preschool children. The appraisal of mental capacity, of course, was made for a variety of purposes, the most frequent of which was diagnosis (81%). Diagnostic testing was reported for such conditions as mental retardation, speech or hearing difficulties, and organic brain damage. Thirty-six percent of the group reported the use of tests as a pre-adoption procedure, and thirty-two percent in connection with admission to nursery schools and other special schools.

Psychotherapy was the most frequently reported type of treatment (52%). Remedial work in relation to such problems as speech and hearing disability, and retardation in reading was reported by 21 percent of the group. Other kinds of treatment services reported were parent counseling (9%), physical and occupational therapy (5%), and developmental guidance (3%).

Sixty percent of the individuals reporting were connected with educational institutions or agencies of one sort or another, almost half of which were nursery schools. Residential schools, day care centers, and kindergartens were represented by an additional 25 percent. Special educational services such as work with various handicapped groups, and with the gifted, were also reported (7%).

The reported research activity of our respondents was for the most part concerned with pathology or other abnormal conditions. Studies of autism and anoxia were reported by 13 percent and 8 percent respectively. Developmental studies, usually longitudinal in nature, however, were next most frequently listed (5%). About two percent of the group also reported studies designed to validate infant tests. Some 24 additional research problems were reported, in each case, by only one or two people. These were such problems as hyperkinesis, perceptual disorders, epilepsy, prematurity, neurological impairment, cognitive development, cerebral palsy, etc.

Types of Cases Handled by the Respondents. For more specific information as to the kinds of case problems dealt with, the respondents were asked to check in a list those types of cases which were common in their clinical experience. These results are given in Table 2. The three most frequently checked case categories were retarded children, normal children and those with behavior problems. Next in order of frequency were brain damage cases, those with verbal disabilities, the physically handicapped or disabled, autistic children, and those manifesting psychotic symptoms. Less frequently handled cases were deaf or blind, those with neurotic symptoms, personality problems, and unusually high mental endowment.

Infant and Preschool-Age Scales in Use. Altogether some 51 different tests and scales were checked or mentioned in response to Item 4 of the questionnaire: "Please check below the tests that are in use in your organization with children under six." Twenty-three, or 45 percent, of these, however, were checked or mentioned by one person only. Over one-third of them, furthermore, were not scales designed to measure mental ability. The most frequently mentioned of these latter were the Vineland Social Maturity Scale, the Rorschach, the Bender-Gestalt and the CAT.

By far the most frequently used mental test, even at these very young (preschool) ages, was, of course, the Stanford-Binet. More than 90 percent of the respondents reported the use of this test. The 15 tests of a strictly mental, or "developmental" nature next most frequently reported, are listed in the order of percentage frequency, in Table 3.

Evaluative Statements

One of the items of the questionnaire was, "which of the above tests (those checked as used in each case) do you consider to be of special value and in what way"?

The Stanford-Binet. As one might expect, the tests most frequently used were mentioned most often as having special value and in a greater variety of ways. For example, 17 different sorts of justifications were offered for the wide use of the Stanford-Binet, such as being interesting to children; being well standardized; the material at lower age levels is well within the scope of young children; gives a mental age score; gives clues to personality development; furnishes much qualitative data; gives significant quantitative data; furnishes a good sampling of behavior reactions; good for the differentiation of verbal and non-verbal abilities, etc.

Many inadequacies of the Stanford-Binet were also mentioned (Questionnaire Item 6). The following were typical: not enough test items at each age level in the preschool years; some objects (test materials) are too small; not enough variety of materials; poor motivational value causing single failures at several levels thus prolonging testing time; seems to give spuriously high IQ's on young children at upper ranges; lower limit not extended far enough; too verbal; some areas of giftedness not tested; some items no longer appropriately placed chronologically; too cumbersome.

The Goodenough Draw-a-Man was described as having a somewhat different pattern of special values from that of the Binet. It was mentioned as being little influenced by cultural opportunities; not too time-consuming; good for the evaluation of perceptual-motor and personality factors; measures social-emotional maturity; excellent for indication of body image; useful with children with language handicaps; is simple and interesting; and has established validity and norms.

The negative comments about the Goodenough are also interesting in view of some of the favorable ones. It was said to have doubtful validity, and that there are discrepancies in norms from age to age. Another complaint was that the qualitative factors, and their significance in the child's reactions, are not sufficiently discussed in the literature.

The Wechsler Intelligence Test for Children (WISC) received 13 different special comments regarding its advantages. It was mentioned as having good validity, as being convenient, easy to administer, reasonably brief. Other comments were that it has interesting content, yields several comparable measures of ability and provides both verbal and non-verbal scores. It differentiates manual, from intellectual dexterity, provides clues to potential, and is a help in giving "developmental guidance" to parents.

On the negative side, WISC brought relatively few comments. Most of those offered had to do with level of difficulty. WISC seems too hard for really young children; seems to depend too much on school experience plus average, or above, intelligence to be useful for below-six ages; instructions are too difficult, too verbal for not-so-bright, not very verbal children; lower limit too high, were the critical comments.

The Gesell Schedule, designed particularly for use with infants, received many favorable comments. Some of these were: it is interesting to the babies; comprehensive and easily administered; demands less direct cooperation from the subject; findings easily communicated; limited demands on the child tend to minimize negativism; simplicity of test materials especially good for developmental guidance of parents; the breakdown of the scale into its components: fine motor, gross motor, verbal and adaptive, is especially useful in the adoptive program for babies.

The "inadequacies and limitations" of Gesell's schedule needs to be regarded in the light of the fact that Gesell never meant for it to be, nor claimed that it was, a scale for measuring intelligence per se. Nevertheless, the "developmental ages" and "developmental quotients" derived from its use are based upon observations of what the baby can do, the only basis upon which mental ability in infancy can be appraised. The criticisms from our respondents were as follows: doesn't evaluate intelligence; insufficient diagnostic precision; not really psychometric; far too subjective; needs updating on language development items; spuriously high IQ's based on "muscles only" in early months.

The Cattell Infant Scale, according to our respondents, has the following advantages: especially good for testing young retarded children; useful for determining approximate developmental level; useful because it runs easily into the Binet, and is therefore valuable in the longitudinal study of mental development; good for evaluating perceptual-motor, and personality factors.

On the negative side, the Cattell is too difficult to score (i.e., too subjective) at some points; its validity is questionable; has only one form, consequently retests are difficult to evaluate; has poor predictive value; its items are too dependent upon physical development.

The Ammons Full Range Picture Vocabulary Test is a relatively new test, published in 1949, yet it ranked in sixth position in terms of frequency of use by our respondents. It also received some favorable comments: it "rules out receptive aphesia" in expressive and receptive aphasic children; especially good with non-verbal and severe physical handicaps; valuable in cases with deficit in language or speech areas; it is a good check of verbal comprehension without requiring verbalization.

The Ammons test was criticized as involving chance errors that are difficult to eliminate; as having norms that are too high, and that progress from easy to difficult concepts is too rapid.

The Merrill-Palmer Scale was praised as "a good ice breaker"; it allows unusual freedom in administration; has a minimum of verbal requirements, and materials and pantomine directions are easily understood by young children; is action oriented; allows flexibility in scoring; taps skills not included in other tests; it shows the degree of planning and ability to carry a task to completion; useful in testing hyperactive, brain-injured children.

Its drawbacks, on the other hand, were seen in its preponderance of performance items. Its verbal portions were found not to be productive; tends to give too high a score; questionable reliability and validity; manual much too difficult to follow and to teach beginners to use; the original manual is out of print; the scale with its equipment is too expensive.

Limitations and Suggestions for Improvement

By way of summary, the most frequently mentioned limitations of the currently used infant and preschool tests when taken together can be grouped under the following headings:

Poor validity

Manuals inadequate or difficult to use

Limited norms

Poor predictability

Insufficient diagnostic precision

Culturally outdated

Give inadequate picture of child's functioning

Too subjective

Lack of theoretical rationale regarding dimensions measured

To the questionnaire item asking for suggestions as to how the available tests could be improved (Item 7) the responses, of course, were many and varied in terms of specifics, but were reducible to a relatively few general ideas. The following summary statements represent fairly completely the individual statements made.

1. Better placement of items, and better manuals and interpretative aids.

2. Need more inclusive predictive materials evaluating parents' educational, social, cultural background, developing ways of discounting or evaluating socioeconomic differences.

3. Have some type of free play situation for assessing emotional status of the young child.

4. Modernize, make less cumbersome, as short as possible, reasonable cost, ease in administration.

5. Should be less subjective, better validation.

6. Need more elaboration, graded series of increasing complexity. Separation of cognitive and perceptual processes. Use of Piaget's findings as theoretical guide for tests of cognitive ability.

7. More data on infant responses and possible meaning as to their potential.

8. More attention to the non-speaking child, the disturbed, less dependence on verbal instructions; behavioral items to detect neurological, emotional aspects.

9. Alternate forms would be helpful.

10. More systematic evaluation of abilities, more reasoning problems of a non-verbal nature.

11. Need new ideas and new approaches.

12. Replace MA concept with a deviation unit scale.

13. A need for factor analytic studies in each of the significant areas.

14. There is a false security in quantification without enough knowledge of the significance of the results.

15. Need for more attention to situational approach in mental functioning.

16. More tests for testing limits, grouping of items into measures of problem solving, motor coordination, memory, abstract judgment, etc.

17. Tests like the WISC should be extended downward and the Merrill-Palmer upward to cover the 5 to 7 year range.

18. More definitive in the developmental areas.

Many of our respondents were kind enough to make additional comments in appended notes or letters. These are greatly appreciated and will prove valuable in further work on the problem. Limited space prevents a complete presentation of this material in the present report.

V

FACTOR ANALYSIS OF INFANT AND PRESCHOOL TEST RESULTS

As was brought out earlier in this report, infant and preschool tests in the past have suffered much disfavor because of failure to predict later mental test results. If an infant test did not agree with later test evaluations it was considered invalid. Several fairly well standardized infant tests have been discarded for this reason. The 1928 Linfert-Hierholzer Baby test, for example, was rejected as completely unreliable because of a low correlation obtained in a longitudinal study comparing scores of infants on that test with later IQ's at the age of 4 years on the 1916 Stanford-Binet. The validity of the infant test was considered shattered in spite of the fact that in standardizing this early Stanford-Binet, only 10 three-year-old children and 51 four-year-olds were used, with only 12 test items, while for the Linfert-Hierholzer, 300 infants were tested, 50 of them repeatedly, with 64 separate test items for the age range of one year.

Now, 30 years later, a need for an evaluation of the infant and preschool test situation is even more in order. Examination of the contents of presently used test scales indicates great variation in content and in breadth of coverage. Some tests, such as the Gesell, cover a wide variety of items, and a wide range of ability is tested at successive age levels. The California

Infant Scale (now the Bayley revision) has separated the gross motor developmental aspects from the more obviously intellectual, making two scales out of content similar to that of the Gesell Developmental Items. On the other hand, the Cattell Infant Intelligence Scale attempts to limit the items to those more comparable to the Stanford-Binet Scale so that it can be a suitable substitute for this scale at the age levels below two years.

Of the two major preschool scales studied in our present analyses of test content, the Stanford-Binet is presumably oriented more toward verbal-intellectual tasks as contrasted with the so-called "performance" type of response. Actually, the Stanford-Binet has several manipulative tasks, such as copying the circle, square, and diamond, doing the three commissions, and discrimination of forms. The Merrill-Palmer Scale, on the other hand, has less language and more object-oriented tasks. As will be shown later, the distinction between these two classes of test items is purely conjectural and has no validity in terms of test item meaning. Perhaps the major significant difference is that the object oriented tasks offer more interest and hence the motivation for responding is greater.

The abilities covered by the currently available infant and preschool tests can be seen, in our analyses herein reported, to be fully as inclusive as those covered by test scales for older children. In most cases there are wider ranges of ability included in the coverage.

The determination of content meaning has not been adequately carried out for many of the present mental test scales, especially at the infant and preschool levels. There is need to determine factor meanings of test items throughout the entire developmental range, since comparability of test results is dependent upon similarity of factor meaning at the different age levels. The development of more refined ways of evaluating test content makes possible such comparison.

It is now seen that our mental test scales contain a variety of test items chosen because they discriminate difficulty levels rather than because of content meaning. The similarity of meaning between two age levels is largely fortuitous and varies greatly from one age level to another. Hence, even with the same test scale, children of different ages are subjected to test items that may have widely different significance. Consequently, the test results for different age levels may not be similar in terms of meaning and the resultant scores for retests are not necessarily comparable.

Since the selection of test items has been haphazard and not systematically determined and because judgments as to meaning of the behavioral response to the items have been highly subjective, the failure to predict later development in terms of some later test score does not necessarily indicate inadequacy of an infant or preschool test. There is no evidence that the results of infant or preschool tests show more variability in longitudinal studies than do mental tests for older age levels. Results of most studies over wide age ranges are remarkably similar in the finding that the percent of children for whom there is complete agreement in scores at two widely different age-levels is usually not more than 10 percent, while 60 percent or over, have IQ variations of 15 percent or more (Ebert and Simmons, 1934; Sontag, Baker and Nelson, 1958). These follow-up studies have aided in dismissing the idea that the IQ is constant over any period of childhood. Honzik, McFarland and Allen (1948), warn that "the observed fluctuations in the scores of individual children indicate the need for utmost caution in the predictive use of a single test score or even two such scores" (p. 315).

The use of different test scales at the various age levels further complicate the process of interpretation of mental growth. In her analysis of the growth of intelligence, using mental test scales suitable for the two divergent age levels, Bayley (1955) found that when her 61 infants had reached 8 years, only a fifth of the group had maintained any stability in their relative status over this eight-year period. Yet the variations found by Bayley in comparing test results from two different scales are no greater than those found in the longitudinal studies of the Stanford-Binet Scale at widely divergent age levels.

The literature on test prediction thus fails to show that infant and preschool tests are less adequate than tests in later childhood in the prediction of intellect at older ages. Neither is there acceptable evidence that they are less valuable in the determination of mental ability than the retest scales with which they are compared. If there were indications that these measures of later-age ability were in themselves reliable predictors, this lack of agreement might be more serious. It can only be assumed in evaluation of the longitudinal findings on mental tests that prediction, to have any merit, must be based on more knowledge about the individual than most mental tests now available can give.

With the advent of more refined statistical procedures for item selection, such as the present day factor analysis methods, we are now able to re-evaluate test item meaning and our concepts of intellect. More and more the evidence tends to indicate the validity of the multiple factor structure of intellect. From such analytical studies of mental tests a much needed revision of our theories of intellectual development can be made.

Comparability of test item content partially determines the potential predictability of mental test results. If two test scales vary widely in content meaning it is unreasonable to expect similarity in total score results. The development of more refined ways of evaluating mental test content makes possible such comparison between test content meaning in different test scales, and at differing age levels.

There is need therefore to determine factor meanings of test items throughout the entire developmental range, since comparability of test results is dependent upon similarity of factor meaning at the different age levels.

Problems Encountered

Factor analysis of mental test material from young children has many pitfalls not always seen by earlier investigators. Earlier results cannot therefore be taken at face value. A total score derived from a combination of several scores that may be measures of several factors, when used in factor analysis, gives ambiguous results because we do not know what combination of the factors may be represented by any individual's score. When the total scores from different tests are used as the basis of the factor analysis the resultant factors tend to coincide with the separate test scales and give little more information about the tests than we had to begin with. This type of factor analysis is illustrated by the one made by Hofstaetter in the Berkeley Growth Study (1954).

The Problem of Experimental Dependencies. It is generally conceded that experimentally dependent variables should not be included in a correlation matrix for factor analysis. Some of the earlier studied, however, have clouded the issue by including total scores (such as MA, IQ or standard scores) for at least one of the test levels as part of the array of variables.

Furthermore, two levels of difficulty for the same test item should not be included in the same factor analysis. Analysis of items for the Merrill-Palmer scale illustrate this complication. Although in scoring this point scale, two or more levels of performance for one test item, such as the Little Pink Tower, were usable as variables in one narrow age-range distribution, yet failure at the lowest level automatically meant failure at the higher age levels, and produced dependent variables that were not valid either for the tetrachoric correlations or for the factor analyses. Our own earlier factor analyses had to be discarded because of this, and other types of errors in selection of items to be included.

The difficulty level increases so rapidly in a series of test items for young children that a very narrow age range is essential to minimize the effect of maturation on their performance. This effect is strikingly present in all the test scales included in this study and is largely due to the rapid rate of development of infants and young children. The condition that might be called a series of irreversible maturity indicators is present, producing a situation where if item A is failed, item B must also be failed. For example, in the Gesell tests (where many such series exist), if Peter cannot sit with slight support, he cannot sit alone; if he cannot regard the pellet, he cannot secure it with his whole hand nor can he show fine prehension of it. In the Merrill-Palmer scale only one child out of more than 300 could copy a star if he failed to copy a cross. While an inspection of the test items may reveal this condition, there are in most scales tests, hidden dependent relationships that are not readily detected by inspection. In the 36 month level, out of 188 cases, no child was able to copy a cross who could not copy a circle, and no child who was unable to close his fist and move his thumb, could pass the more difficult test of touching the tips of his fingers with his thumb. This, of course, makes it necessary to eliminate one of the two items in each case.

Thus it is clear that the selection of variables for factor analysis needs careful scrutiny. While one wishes to include as many variables as possible in any matrix, nevertheless, some items that, on first inspection, seem to be usable should be eliminated.

It is possible that the reason tests at younger age levels do not present a reliable prediction of later age mental test ability is that the tests at the younger age levels are differently weighted with a wider variety of factor content, or they may have a smaller number of factors than are included in older age levels. The number of usable factors extracted is dependent upon the number of variables included in the factor analysis. Hence, there is no necessary developmental significance to obtaining a large or small number of factors. The number of items that should be used in any one age-range distribution is dependent upon the variety and breadth of item meaning included in the sample. A wide variety of items is essential to obtain a number of different factors.

Most test scales have items that recur with increasing difficulty as the age levels of the scale increase. In selecting the items to be used in the factor analysis of one particular age level, the items most characteristic of that age level should be chosen, eliminating the rest. There should be no duplication, no dependent test items. The necessary reduction in the number of test items in this study has been done with some reluctance, since the difficulty level itself has limited the number of usable test items. The range of items included within the limits of 90 percent success and 90 percent failure is usually rather narrow for any one defined age level.

Since the usual test items at the infant and preschool levels reflect a very rapid process of maturation, in any test scale the number of usable items for analysis is very much curtailed. For example, in our factor analysis of the 54 month level of the Merrill-Palmer scale, where the distribution is limited to only one month variation in age range, and number of usable items from the point of view of percentage passing was 27, but when all dependent variables were eliminated, only 11 items remained in the correlation matrix.

The method of identifying hidden dependencies involved a tabulation of the b and c values for the tetrachoric table used in obtaining the correlations for each set of variables. From these computer-derived tables for each pair of items, the probability was computed of passing item y if the less difficult item x had been failed. Thus, the relative difficulty of items could be derived as well as the probability of passing y for each pair of items in the correlation matrix. Whenever it was shown that failure on item x gave a probability of success of less than .10 on the more difficult item y, it was conceded that y was, in some way, dependent upon the abilities involved in the less difficult item, and one of the two items was eliminated.* Using this procedure in the selection of test items for analysis, the number of communalities greater than 1.00 in the resulting table of factor loadings was reduced to a minimum compatible for an acceptable factor analysis.

---

*A discussion of the findings obtained in this process of elimination of dependent variables will be presented in a separate paper.

Determination of the Number of Factors to be Extracted. There has also been some confusion in interpretation as to the significance of how many factors can be extracted from a correlation matrix. In the earlier factor studies it has been observed that the number of isolatable intellectual factors decreases with decreasing age, and an inference was made that the test performance and capacities of younger age groups, particularly at the preschool and infant levels, were increasingly general and less definable in terms of the specific factors discernable at the older ages. As will be shown in the present study, modern electronic computer techniques for factor analysis and factor rotation make possible a more complete factor analytic approach to the determination of item meaning. Since the number of principal axes obtained is never greater than the original number of variables used, the smaller numbers of factors obtained in our final factor tables are due merely to a reduced number of test variables.

The number of items included in each of our correlation matrices, as we have indicated, has been much reduced from the original list. Every factor matrix in the tables to follow is, therefore, the result of the discovery of hidden dependencies in previous, discarded factor analyses. Fortunately, the ease with which computer factor analyses are done, made possible the continuous elimination of dependent variables by repeated factorizations until satisfactory item selection had been made. On the CDC 1604 computer used in this study, the required computer time for performing one factorization is about 30 seconds from a 25 x 25 matrix.

Factor Identification and Interpretation. Perhaps our biggest challenge in the factor analysis of infant and preschool test items is that of factor identification. The interpretation of factors is necessarily affected by the factor methods used and the variables included in the correlation matrix. As we have already indicated, early studies that were done by laborious hand methods of extraction included fewer factors.

In the naming of factors there has been in the past a tendency to label any test item involving manipulatable test material as a performance test in contrast to those items where the content is purely verbal. These latter were labelled "intellectual," "cognitive," "reasoning," or "memory" tests. The early factor names reflected this pattern. Surely the naming of factors has had much alteration since the time when these early factor studies were made. Earlier theories about the changing nature of intelligence from infant to school-age level have been influenced greatly by the factor naming in these early factor studies.

The naming of factors is facilitated by:

(1) inclusion of test items of generally accepted factor meaning (marker test items),

(2) familiarity with the item content of the test scale,

(3) realization that the name given to the test item usually refers to the type of test materials used and does not describe necessarily the mental processes involved,

(4) ability to analyze the behavioral responses and processes involved in carrying out the task set by the item,

(5) avoidance of mental bias resulting from conjectures and theories about mental development, and

(6) realization that the meaning for any one item may change from one age level to the next.

There has been an attempt in this study to offer names for most of the factors obtained. These names are not given authoritatively but only as tentative suggestions, for much corroborative study needs to be done to determine definitively the correct meaning of different factor groupings of individual items.

General Procedure

Test protocols for infant and preschool tests were available from many sources. It was decided that a thorough and careful factor analysis study of the item content of these tests might be of value in understanding the similarity of factor meaning at different age levels and for different test scales. Consequently, test records were tabulated by items for five different infant and preschool scales. Table 4 gives the tests used, with the number of children for each of the 14 age-level groups studied. The combined data for 1,926 children gives a continuous range of developmental progress in intellectual ability from 3 months to 5 years.

Statistical Procedure. From each of the fourteen age level distributions, inter-item tetrachoric r's were computed from the cosine-phi approximation. The principal axes were extracted and rotated to the orthogonal varimax criterion (Kaiser). In each case factorization was iterated until all communalities were constant to within .02. Computations were made with a CDC 1604 computer at The University of California at San Diego.

Several experimental studies were carried out to discover which of two methods was better, for determination of the final communalities for material of this kind. The first method was to take as the trial communality in each case, the largest element in the row of the factor table. The second was to set the initial value equal to one. In each case, the factorization was repeated until all communalities were constant to within .02. The final results were very similar, indicating convergence in spite of the starting value. This method of iteration was used in computing all factor values. It is a method that is unique and independent of the choice of the initial communalities.

In the original selection of items to be retained, the criterion was set at 90 to 10 percent passing or failing, for the tetrachoric correlations. However, as we have seen, it was found later that with young children this standard was too inclusive, and further extensive cuts were necessary. Thus, it was necessary to eliminate many of the test items we wanted to retain.

However, when these reduced lists of items were processed by the computer, and principal axes extractions and the varimax orthogonal rotations were made, it was found that some of the resultant rotations were still far from a satisfactory interpretation, due, in part, to the unavoidable fact that the reduced sets of items did not provide enough measures for each kind of ability to "uniquely determine" the factors, as varimax seems to require. When we use varimax we must estimate very carefully how many factors we want, for if we choose too many, it will tend to split up the factors and if we use too few we will obtain big clumps that often yield a semblance of a general factor, or gross group factors. Varimax tries to give each of the variables as much weight as possible and therefore tends to split up meaningful combinations.

Since, at this point in our exploration, there was sufficient evidence that several kinds of intellectual functioning can be differentiated in the very young child, we proceeded to try further rotational methods. Because of this difficulty with the computer-produced varimax, graphic orthogonal rotations have been introduced for each of the factor analyses in order to "tidy-up" the confusion evident in the varimax rotations.*

In addition, from the varimax results, oblique rotations have been worked out by computer, using the "biquartimin" criterion, (with parameter gamma = .5). Biquartimin has the advantage of lower correlations between the primary factors, and thus more nearly approaches orthogonality if the structure is really orthogonal. For each of the rotations the "normalized loadings" of the initial factor matrix were used.

The complete set of results from each of the 14 final analyses, including the correlation table, the principal axes factor matrix, and the three sets of rotation results, is presented in a single table (Tables 5 through 18).

Interpretation Procedures. In the naming of factors we have used the "operations" included in Guilford's structure of Intellect, in order to see what contribution they make to our understanding of the nature of mental ability at the lower levels of human development. While the naming of factors is necessarily of a somewhat subjective, intuitional nature, nevertheless it is possible to line up most of them tentatively with known factor content at older age levels, as can be seen by the data presented below. The conjectural nature of factor names is reduced by a careful analysis of the various functioning processes, and by having the factors named independently by several qualified persons.*

---

*We are indebted to Philip R. Merrifield for carrying out all our graphic rotations.

Guilford's structure of Intellect (1960) involves five intellectual operations: Cognition (C), Memory (M), Divergent Production (D), Convergent Production (N), and Evaluation (E). These five processes can each be classified into four general types of information or content. Each of these may be divided again into the results of the individual's processing of information.

Guilford's classification system can be seen roughly to apply to the test performance of young children. However, the expressions of these abilities do change in a qualitative sense with age. Furthermore, most of the earlier analyses of these factors were made on adults and adolescents, which makes the identification of the sub categories much more difficult at the preschool level, particularly since the structure of early mentality is likely to be relatively less complicated.

---

*We have had the privilege of having Guilford's advice and assistance in this factor naming, and we also have had much help in this process from Merrifield who has worked extensively with Guilford.

In the interpretations to follow, the Guilford model is used whenever the items constituting the factor made meaningful such an interpretation.

The variously rotated factors for each age level are interpreted in the light of the known item content, and, as with most factor naming, there is much uncertainty as to the basic psychological processes involved. While a limited statement of the nature of the test item is presented here, the reader is referred to the original description of items in the test manuals for greater detail.

In order to facilitate consistent interpretation of factors with a psychomotor meaning, Guilford's System of Psychomotor Abilities (1958) is used.

The Choice of a Rotation for Factor Interpretation. In the pages that follow we discuss the results of the separate analysis at each age level for each of the five test scales studied. In each case, a comparison of the factor loadings from the three rotational procedures is given. Results of the graphic rotations are assigned letter notations: A, B, C, etc., to distinguish them from the two machine rotations. In certain cases there was lack of agreement between the rotations in their distributions of highly loaded items among the various factors. In most instances however, it was found that the graphic rotation gave the most easily interpreted results, and in consequence was chosen consistently for the age-level factor interpretations.

The Cattell Infant Intelligence Test

One of the most frequently used mental tests for infants is the Cattell Infant Intelligence Test. The age range on this scale extends from neonate, well up into the preschool levels. It was designed to be a supplement to the Stanford-Binet scales at age levels below those covered by that scale.

In our selection of cases for this scale we had hoped to have age-level distributions at three months, six months and twelve months. Unfortunately, the different sources yielding an abundant number of tests for the younger ages, were able to supply only 52 protocols at age-level 12 months, and these were under suspicion as not representing a normal sampling. Hence our data were limited to the three and the six months age-levels.*

A particularly narrow range of ability was tested by the two Cattell age levels. The inter-item correlations were all very high.

The Three-Months Level. All the statistical tables for the three months level of the Cattell scale are combined in Table 5. Initially there were 13 items in the distribution, but only 8 were retained because of dependent relationships. The ability range is very narrow and definitive, and items have much similarity in content. Over 90 percent of the 3 month-old children passed all the items at the 2 months level, and less than 10 percent passed any one of the items at the 6 month level. For most of the 5 month items the probability of failure was higher than .90. All but one of the 28 independent correlations in the matrix was over .50. The inter-factor correlations (Table 19) for the biquartimin rotations for the two Cattell age-levels were higher than for most of the other distributions, but not as high as one might predict from the high item inter-correlations obtained. The varimax solution shows no hyperplane and cannot really be interpreted. The graphic, admitting one facilitating factor (A) has fair hyperplanes for B and C.

---

*The chief sources for these protocols were from the early records of The Merrill-Palmer Institute, Libby Bass (St. Louis) and Marie Skodak (Dearborn).

Factor loading over .40 follow for the various factor rotations giving item number in rank order:

Varimax			Graphic			Biquartimin		
1.	3	.884	A	3	.96	1.	3	.999
	1	.776		4	.95		1	.779
	4	.758		1	.94		4	.751
	5	.563		5	.83		5	.422
	8	.505		8	.79			
	2	.451		2	.74			
				6	.65			
				7	.61			
2.	7	-.814	B	7	.70	2.	7	-.887
	8	-.645		6	.56		8	-.579
	6	-.569		8	.45		6	-.411
	4	-.475						
3.	6	-.705	C	6	.49	3.	6	.629
	5	-.679		5	.45		5	.582
4.	2	.612	D	2	.41	4.	2	-.545
	1	.453						

The three factor rotations show a fair equivalence in the first factor loadings at the 3-months level. An inspection of the graphically rotated factor (A) would seem to indicate a general factor.* The varimax places two of the items below a minimum of .40 loading. The biquartimin seems to yield the most simple structure with easily identifiable range of factor loadings, are as follows:

Item	Range of Factor Loading
3. Regards cube	.999 to .884
1. Following ring	.95 to .779
4. Regards spoon	.94 to .751
5. Inspects fingers	.83 to .422
8. Increases activity at sight of toy	.79 to .505
2. Anticipates feeding	.74 to .451
6. Finger manipulation	.65
7. Hands remain open	.61

In the biquartimin set of items there is no motor response, other than the eye fixation, involved. With all the rotations it can be seen that the heavily loaded items require visual inspection or visual response to an object. The suggested name for this factor is Attention of a Visual type - visual cognition, (CFU).

The three high loadings for Factor 2 of this analysis are the same for all three rotational procedures. It would seem that a hand response, directed toward an object to which the infant is attending, is the central behavior. High loadings for all three rotations are given:

Item	Range of Factor Loading
7. Hand remain open	-.887 to .70
8. Increases activity at sight of toy	-.645 to .56
6. Finger manipulation	.569 to .411

---

*Graphic Rotations for Cattell age distributions were done by C. Victor Bunderson and Stephen W. Brown (1962).

Suggested Name: Convergent production, hand-arm response to a visual stimulus, (NFR).

Again with Factor 3, the three rotations agree as to factor content. There is a combination of visual and a motor response of the fingers, a kind of awareness of finger functioning. The high loading ranges for three rotations follow:

Item	Range of Factor Loading
6. Finger manipulation	-.705 to .49
5. Inspects fingers	-.679 to .45

Suggested name: Divergent production with fingers, figural fluency, (DFU).

Both the graphic and the biquartimin rotations give only one high loading to the fourth factor but all three rotational procedures are in agreement that Item 2 is most highly loaded. Varimax includes one other item that shows an interpretable degree of relatedness. Both involve adjustment of the eyes to a moving object. Anticipation also involves memory. The summary of high loadings for this factor follows:

Item	Range of Factor Loading
2. Anticipates feeding	.612 to .41
1. Follows ring in circular motion	.453 (Varimax)

Suggested name: Perceptual foresight - cognition, (CFI).

Cattell Scale, Six Months Level. The six months level of the Cattell Scale gives very much the same picture as the three months level (See Table 6). The variety of test items is limited in comparison with the Gesell or California tests for this level. From the original 16 items tabulated at this age level only 8 were usable for the intercorrelations and factor analysis. There were no values under .50 in the 56 correlations and only 4 of these were less than .60. It is small wonder that the factor loadings tend to be very high. The graphic rotation includes all of the items in the first factor. The varimax and biquartimin agree with the graphic rotation on the first three loadings, but exclude the last four which are represented instead in factor D in the varimax rotation. The summary of the high factor loadings follows:

Varimax			Graphic			Biquartimin		
1.	4	.845	A.	4	.98	1.	4	.929
	2	.547		2	.84		3	.448
	3	.525		3	.78		2	.421
	5	.478		5	.78			
				1	.71			
				6	.67			
				7	.67			
				8	.58			
2.	8	-.773	B.	8	.68	2.	8	.881
	7	-.726		7	.48		7	.844
	3	-.544		6	.45		3	.483
	6	-.521		1	.45			
	2	-.426						
3.	5	.740	C.	7	.43	3.	5	.699
	6	.569		5	.42		6	.466
	7	.479						
4.	1	-.734	D.	1	.41	4.	1	-.700
	2	-.511						
	6	-.492						
	8	-.403						

As one examines the first factor (A) of the Graphic Rotation, the narrow range of abilities makes it quite plausible that these eight items all belong in one general factor. However, the factor meaning is not necessarily "general intelligence," for all items have to do with reaching. The responses seem to be those involving reaching out toward some object. It is logical not to label a behavior as "reaching" unless there is an object seen and responded to. The reaching out tends to be with one hand. Item 4 has by far the highest loading and this is defined as unilateral reaching. It can be seen that both varimax and biquartimin cut out the factor loadings for the last four items. This strengthens the supposition that the emphasis in this factor is unilateral reaching. However, to be able to reach also means to be able to perceive objects as attractive, worthwhile attaining. Usually the objective of such reaching is further exploitation and exploration. Hence it should not be immediately assumed that this response is merely "motor." It involves having learned the eye-hand coordination necessary to attain the object and the capacity to perceive its desirability. While the last four items in the graphic rotations are related and involve this reaching complex to some extent it may well be that the differentiation shown by the varimax may yield cues for discriminating. The Graphic factor A is reproduced below for the sake of consistency.

Item	Factor Loading
4. Reaches unilaterally	.98
2. Secures cube on sight	.84
3. Fingers reflection in mirror	.78
5. Reaches persistently	.78
1. Transfers object from hand to hand	.71
6. Attempts to pick up pellet	.67
7. Inspects ring	.67
8. Exploits paper	.58

Suggested factor name: Cognitive responses emphasizing unilateral reaching for an objective stimulus (CFU).

One can only conjecture what the content meaning of the second factor involves. It may be a matter of reaching to feel the object. In each case the child is given an object to manipulate and inspect. The major interest is in exploration. The graphic solution is presented below for determining the factor name.

Item	Factor Loading
8. Exploits paper	.68
7. Inspects ring	.48
6. Attempts to pick up pellet	.45
1. Transfers object hand to hand	.45

Suggested factor name: Exploration and investigation of objects by manipulation. Experiential evaluation, (EMS).

There is fairly good agreement between the three rotations on the third factor with a slight difference in the graphic rotation. The factor seems to be emphasizing persistent behavior, repeated performance in order to obtain an objective. The varimax rotation with its three items emphasizes this behavior slightly more than do the biquartimin and graphic, which omit either item 6 or item 7 from their loadings of .40 or more. Persistent reaching certainly fits this idea, since the reaching in this case must be effortful attempts to obtain the object. When a child attempts to pick up a pellet, repeated action is implied with some failures in the process. Inspection implies looking at an object from more than one position and for repeated visual explorations. The high loadings for graphic factor C are given:

Item	Factor Loading
7. Inspects ring	.43
5. Reaches persistently	.42

Suggested factor name: Persistent behavior toward a goal, *a kind of convergent production of figural units*, (NFU).

All three rotations select the most heavily loaded item in the *fourth factor*, "transfers object from hand to hand," but only the varimax includes any other item with a loading over .40. It is probable that the behavior indicated by this factor is that of manipulating an object. Inspection of the ring may involve moving the ring around to different positions, as well as does the exploiting of paper indicate some varied manipulation of this object. The one significant loading of .41 from the graphic rotation, was for item 1, "transfers object from hand to hand."

Suggested name: *Divergent production* - exploratory manipulation, (DFT).

California First Year Scale

It is remarkable that the California Test Scale developed at such an early stage of test construction should offer so rich a variation in test meaning. Even so, Bayley (1933) placed the gross body response items, and more specifically "motor" items, in a separate scale. The difficulty level determined the order of these test items in a point scale. Our later work with them agrees very well with this early item placement.

One defect in this present analysis is the necessity of eliminating a large number of meaningful test items because the rapid development of the infant establishes a narrow difficulty range. Actually in this scale covering a period of one year, over 100 test items are introduced, but the percentage of success at any given narrow age range, changes rapidly because of the developmental similarity of the group of normal infants. Since our two distributions for this scale covered an age range of just two months each, many of the test items in the total series are eliminated. In order to cover the entire range of items several other age level distributions would be needed.

Bayley contributed 63 test records to our distribution for each of the six and twelve month levels. These added to our original seventy-four made 137 records to be included in our analysis. Bayley has also kindly cooperated, completely independently, in the attempts at naming the factors and identifying the meaning of the factor contribution of each item for the two levels of the California First Year Test that have been factor analyzed. However, in some instances the final factor content was not quite the same as when she gave the suggested name. Bayley made no attempt to suggest names within the conceptual framework of Guilford's Structure of Intellect.

Six Months Level. Of the originally tabulated 19 test items that had been selected on the basis of including all items passed by the middle 80 percent (eliminating the upper and lower 10 percent) only eleven were retained for our final factor study. Several of these 19 items were found to have mediative relationships that necessitated the selection of only one of the inter-related set.

The principal axes analysis and the various rotations are presented in Table 7. The graphic rotation was chosen as most meaningful in assigning factor names for this factor matrix. A summary for the various rotations of the items having factor loadings of .40 or more follows:

Varimax			Graphic Orthogonal			Biquartimin		
1.	4	-.921	A.	4	.62	1.	4	.877
	8	-.414		10	.62			
	3	-.400		8	.42			
2.	11	.890	B.	11	.75	2.	11	-.890
	3	.398		3	.57			
3.	2	-.832	D.	2	.80	3.	2	.862
	5	-.530		11	.46		5	.500
				5	.40			
4.	3	.756	C.	9	.80	4.	3	.733
	1	.679		6	.78		1	.715
	9	.488		8	.68			
				3	.67			
5.	10	-.786		10	.62	5.	10	.813
	8	-.683		1	.61		8	.694
	6	-.587		7	.51		6	.528
	9	-.555		5	.46		9	.463
6.	7	-.795	E.	4	.70	6.	7	-.763
	6	-.448		7	.46			

In the first factor, the three rotations agree in placing Item 4, with the highest loading, but the varimax and the graphic rotations agree on one more item loading over .40, which helps in setting up a tentative factor name. The following are the loadings in graphic rotation (A):

Item	Factor Loading
4. Turns after spoon	.62
10. Rotates wrist	.62
8. Lifts cup with the handle	.42

This factor seems to involve bodily adjustment to an object visually perceived. Wrist rotation is performed in relation to the object that is to be manipulated, but the turning after a fallen spoon would eliminate the manipulation of the object itself but would rather select out the body adjustment, which could also be included in the reaching for the cup by the handle, since this includes rotation of the wrist and hand adjustment in order to attain the handle of the cup. The action is not mere coordination but involves goal direction, and consequently might be considered a cognitive response involving body adjustment to gain a specific goal.

Suggested factor name: Visual cognition of figural relations. (CFU). Bayley's suggested name is, Goal directed attention.

There is agreement between the three rotations that the major significance of Factor 2 is represented by Item 11, but the graphic rotation also includes the lesser weighted item 3, which helps define the factor meaning.

Item	Factor Loadings
11. Smiles at mirror image	.75
3. Reaches persistently	.57

Smiling at a mirror image at 6 months does not mean that the child knows who the person is in the mirror reflection, but he does see a person. Also the reaching persistently implies a person who is stimulating the reaching response by manipulating the object. In both instances there is behavioral content of a cognitive-type response to a person, rather than to an object.

Suggested factor name: Cognition of Behavioral Units, (CBU). Bayley also feels that these items have a social response element.

The graphic rotation combines the items in Factor 4 and Factor 5 of the varimax and biquartimin to constitute the third factor (C). Yet on the whole the combination of variables included in this graphic rotation seems to offer more meaning than do the other factor rotations.

The items over .40 in graphic Factor C are as follows:

Item	Factor Loading
9. Exploitive string play	.80
6. Sustained inspection of ring	.78
8. Lifts cup with the handle	.68
3. Reaches persistently	.67
10. Rotates wrist	.62
1. Accepts second cube	.61
7. Vocalizes satisfaction	.51
5. Bangs in play	.46

All of these items have a similar characteristic, namely, that the response to be given is not inherent in the item itself as presented, but rather allows a kind of selective behavior to result from the stimulus presented. The behavior involved is exploratory manipulation of an object and there is permissible variation in the response. Such behavior as persistent reaching, exploitive string play, and banging in play, stand out as divergent types of responsiveness.

Suggested factor name: Divergent production of figural relations, (DFR). Bayley suggests: Alertness of an exploratory type.

All three rotations show similar factor content for the fourth factor, graphic (D). There is a semantic content to these items, a pre-verbal or pre-language factor with a behavioral aspect. The major loading emphasis is upon vocalizing displeasure. Bangs in play involves strongly an interest in sound production, a responsiveness to the environmental stimulus situation. There is evidence of evaluation in the responses. The graphic orthogonal rotation (D) is as follows:

Item	Factor Loading
2. Vocalizes displeasure	.80
11. Smiles at mirror image	.46
5. Bangs in play	.40

Suggested factor name: Evaluation of Semantic Systems, (EMS). Bayley suggests: Extrovertive responses not goal centered.

The last factor in each of the three rotations gives item 7, vocalizes satisfaction, a high loading There is no agreement on the other items. This factor, E of the graphic rotation, is as follows:

Item	Factor Loading
4. Turns after fallen spoon	.70
7. Vocalizes satisfaction	.46

This would seem to be a factor offering some semantic affective quality, but with an evaluative implication. There seems to be little indication that this factor is merely a residual one. The suggested name for this factor is Evaluation of Semantic relations, (EMR).

Twelve Months Level. The original tabulation of 31 test items used in our early factor analyses was reduced to 13 because of experimental dependencies. The three rotations of the 12 month principal axes showed much agreement. The biquartimin and the graphic orthogonal have reduced slightly the number of high loadings from those found in the original varimax rotation, and the graphic rotation has cleared up some of the fuzziness of the other two for the lesser factors. The complete results of this analysis are given in Table 8. The comparative patterns of the three factor rotations follows:

Varimax			Graphic			Biquartimin		
1.	2	.886	A.	1	.79	1.	2	.921
	1	.781		2	.76		1	.788
	4	.761		4	.71		4	.771
	3	.575		5	.56		5	.531
	5	.525		9	.51		3	.527
	8	.500		3	.48			
	9	.426						
2.	11	.809	B.	11	.80	2.	11	.882
	13	.715		13	.61		13	.604
	8	.526		8	.54		8	.537
3.	6	.655	C.	13	.61	3.	5	.611
	5	.636		6	.64		6	.682
	13	.467		5	.53		13	.428
4.	12	.663	F.	12	.58	4.	12	.712
	9	.516		9	.51		9	.476
	3	.450		3	.43		3	.413
5.	10	.636	D.	8	.68	5.	10	.619
				2	.47			
				10	.44			
6.	7	.786	E.	7	.78	6.	7	.803
				10	.54			

The graphic rotation (A) seems to give the most unified picture of factor content for this first factor.

Item	Factor Loading
1. Attempts to secure three blocks	.79
2. Picks up cup, secures cube	.76
4. Unwraps cube	.71
5. Holds crayon adaptively	.56
9. Stirs with spoon in imitation	.51
3. Puts cube in cup on command	.48

While one might be tempted to call this "cube behavior," more careful scrutiny gives insight into the possible processes involved. The scoring of these items is not based upon the degree of skill shown, but upon whether the baby can do the task because he understands what to do. The examiner is present with directions as to what should be done in each case, and therefore comprehension of both verbal and demonstrated directions is involved. Suggested factor name: Deduction, convergent production of semantic implications, (NMI). Bayley proposes: Concepts of relationships.

There is complete agreement on Factor 2 as to the content meaning. This is an early language factor.

Item	Factor Loadings		
	Varimax	Graphic	Biquartimin
11. Says two words	2. .812	B. .80	2. .882
13. Uses expressive jargon	.702	.61	.604
8. Imitates words	.530	.54	.537

Bayley suggests that this is a communication factor. The proposed name for this factor according to the Guilford categories would not be semantic, but rather: Convergent production of figural units (NFU).

Again the three rotations are in agreement as to the high loadings in the third factor, placing Item 6, "Inhibits on command," first in order of importance.

Factor C of the graphic rotation is reproduced here for reference.

Item	Factor Loading
6. Inhibits on command	.64
13. Uses expressive jargon	.61
5. Holds crayon adaptively	.53

This factor seems to involve an adaptive response to the directions of the examiner. The use of expressive jargon often accompanies pointing, in an attempt to express verbally the baby's reaction to the situation. The behavior here seems to be cognitive in nature. We have named it tentatively: Cognition of behavioral units, (CBU). Bayley's suggestion: Goal directed behavior.

Both varimax and biquartimin agree on the first three loadings in this fourth factor. The varimax rotation, however, adds item 8, imitates words. The graphic orthogonal rotation for this factor (F) agrees with the other two rotations. The graphic rotation gives the following high values.

Item	Factor Loading
12. Dangles ring by the string	.58
9. Stirs with spoon in imitation	.51
3. Puts cube in cup on command	.43

Item 12 involves a reproduction of the demonstrated dangling of the ring by the string. The other items also require an imitative response.

Suggested structure of intellect category: Cognition of figural relations, (CFR). Bayley suggests: Immediate memory.

The graphic rotation divides the loading for Item 10, Holds cup to drink from, between the fifth and sixth factors (D and E), while the varimax and biquartimin rotations give it uniquely high loadings in factor 5. Factor E in the graphic rotation also includes two highly loaded items, in contrast to the single item 7, Strikes doll in imitation, in the corresponding factor 6 of the other two rotations. The graphic orthogonal rotation for both of these factors (D and E) seems to be preferable.

Factor D

Item	Factor Loading
8. Imitates words	.68
2. Picks up cup, secures cube	.47
10. Holds cup to drink from	.44

There is a kind of imitative memory implied in these items. The suggested factor name is: Memory for figural relations, (MFR).

Factor E

Item	Factor Loading
7. Strikes doll in imitation	.78
10. Holds cup to drink from	.54

There is a social imitative aspect to these two items. They are behaviors related to the cultural experiences of the child. Bayley suggests that these items involve the repetition of earlier experiences that have been "taught." The suggested factor name is: Memory for behavioral units, (MBR).

Gesell Developmental Schedule

The Gesell developmental items are taken from specific age levels. The scoring was done on the basis of pass or fail for each item at the two age levels, six months and twelve months. All the records were taken from the files of the Merrill-Palmer Institute, some from the intensive case records and many from the infant study files.

Gesell, 6 Months Level. The six month level of the Gesell Schedule involves a large variety of items. Many of these were eliminated as they were copied from the original protocols because they were obviously so much like other items that the dependency relation could be detected by inspection. From our original list of 42 items there were still many other items that proved to be experimentally redundant. With further reduction, eliminating those having duplicate significance, there were 19 items remaining. See Table 9 for the complete analysis. There were 8 factors for both varimax and the graphic rotations. A biquartimin rotation was not made for this age level. A summary of factor loadings of .40 or greater follows:

Varimax			Graphic Orthogonal		
1.	5	.826	A.	11	.77
	11	.813		12	.67
	12	.808		5	.59
	16	.546		16	.49
	15	.518		7	.43
	3	.496		3	.42
	7	.427		15	.42
2.	13	.957	G.	13	.98
	10	.459		10	.51
	18	.425		18	.44
3.	7	.856	C.	7	.71
	17	.770		8	.71
	8	.665		17	.66
	16	.523		19	.64
	19	.444		16	.53
4.	2	.937	D.	2	.42
	8	.408		6	.40
5.	9	-.890	E.	9	.83
	3	-.396			

6.	4	-.908	F.	4	.87
	6	-.767		6	.75
	19	.411		7	.54
7.	14	.814	B.	14	.80
	18	.559		15	.54
	15	.511		10	.53
	10	.476		18	.46
				5	.42
				8	.40
8.	1	.828	H.	2	.76
				1	.65
				18	.40

With slight variation in order, factor A of the graphic presentation coincides with factor 1 of the varimax rotation. The item content of the heavily loaded items of graphic factor A follows:

	Item	Factor Loading
11.	Picks up cube	.77
12.	Lifts inverted cup by the handle	.67
5.	Manipulates spoon, etc.	.59
16.	Crumples paper	.49
7.	Reaches directly	.43
3.	Makes stepping movements	.42
15.	Reaches above head for dangling ring	.42

Each of these items requires capability of manual closure. The production is one involving reactions to things to be manipulated in exploratory ways. The manner in which they are handled has not been prescribed. The suggested factor name: Divergent production of figural relations, (DFR).

The tests with high loadings on varimax factor 2 and on graphic factor G are identical. These items follow with their graphic factor (G) loadings:

Item	Factor Loading
13. Looks for fallen object	.98
10. Regards pellet	.51
18. Expresses recognition	.44

Memory is involved quite obviously in items 13 and 18, and, while it is difficult to say what mental reaction is occuring when the baby is regarding the pellet, cognition certainly cannot be ruled out as being involved. However, the memory aspect seems more significantly related. Suggested factor name: Meaningful Memory, (MMR).

Factor 3 and factor C include the same high loadings in slightly different order. The graphic rotation placement is reproduced here:

Item	Factor Loading
7. Reaches directly	.71
8. Manipulates spoon with one hand	.71
17. Purposeful reaction to paper	.66
19. Splashes in the tub	.64
16. Crumples paper	.53

Here we find the reaching directly item, which must involve an object to be reached, and the manipulation of a spoon with one hand, which is a visual response of an exploratory type, both of the same high-loading value. The paper reaction is also exploration of two types (items 17 and 16). It seems that these items might all be said to involve visual exploration of the environment. Suggested factor name: Evaluation of semantic systems, or Guilford's name, "experiential evaluation," (EMS).

The fourth factor, graphic (D) has two acceptable loadings, but neither of them is high enough for much interpretive significance.

They are:

Item	Factor Loading
2. Creeps or hitches	.42
6. Bangs spoon	.40

The nature of these items, however, suggests that we have have here a factor that selects out *locomotor ability*. This interpretation is supported by the very high loading of .937 for item 2, *creeps and hitches*, in the varimax rotation.

Both rotations give the *fifth factor* very high loadings for item 9, *blinks at pencil*. The loadings for item 3, *makes stepping movements*, are slightly lower than our minimum of .40 in both instances but they may give us clues as to the factor meaning here. Both are reflexes that indicate increasing neural development. Suggested name: *Reflex development*.

The two higher loadings of varimax *factor 6* agree with those of graphic factor F. The graphic rotation gave results as follows:

Item	Factor Loading
4. Pats table	.87
6. Bangs spoon	.75
7. Reaches directly	.54

The varimax rotation gave a different item for the third high loading (No. 19, *splashes in the tub*). This activity certainly involves noise making, and, as do the first two, it also involves rapid flexing of the arms. However, there is also interest in exploiting the substances and materials in the environment that are novel to him. A suggested name for this factor: *Divergent production of figural units*, (DFU).

The high loadings of factor 7 (varimax) are similar to those of graphic factor B. The graphic (B) loadings are as follows:

Item	Factor Loading
14. Holds two cubes	.80
15. Reaches above head for dangling ring	.54
10. Regards pellet	.53
18. Expresses recognition	.46
5. Manipulates spoon	.42
8. Manipulates spoon with 1 hand	.40

These items all involved production of a visual type to a stimulus provided by the examiner or other adults in the environment. Suggested factor name: Convergent production of figural relation, (NFR).

The eight factor of this analysis has clearest meaning in the graphic factor H:

Item	Factor Loading
2. Creeps or hitches	.76
1. Rolls from back to stomach	.65
18. Expresses recognition	.40

The two highest loadings here are more suggestive of whole body movements -- a whole-body factor. Expressing recognition may involve a mass reaction response also, but this is not so significantly indicated. Suggested factor name: Psychomotor whole-body control.

Gesell, 12 Months Level. The varimax and biquartimin rotations in this analysis agree very closely, both in size and content of factor structure for all the factors. The graphic pattern however, includes fewer factors, each generally containing more items. In some instances this makes for ease in factor definition, at other times the interpretation is less clearly evident.

The 35 originally selected items were reduced in number to 19 because of dependent relations that were not always evident by inspection. Even with this reduction, eight clearly distinguishable factors were interpretable with both the varimax and biquartimin rotations. Only six were interpretable in the graphic orthogonal rotations. Table 10 gives the complete factor analysis data. A summary of the three kinds of rotations follows:

Varimax			Graphic Orthogonal			Biquartimin		
1.	8	.789	C.	7	.70	1.	8	-.884
	18	.749		8	.62		18	-.848
	9	.615		9	.52		9	-.640
	2	.512		6	.52		2	-.479
				2	.48			
				18	.42			
				12	.41			
2.	3	.984	B.	3	.91	2.	3	-1.015
	15	.689		15	.70		15	- .716
				6	.46			
				17	.44			
3.	5	-.753	E.	5	.72	3.	5	.796
	11	-.711		19	.65		11	.771
				6	.56			
				11	.43			
4.	4	-.799	D.	2	.62	4.	4	.805
	1	-.496		1	.61		1	.460
	7	-.434		18	.58			
				19	.56			
5.	19	-.904		4	.45	5.	19	.859
				16	.43			
6.	12	.856	F.	12	.66	6.	12	-.900
	2	.526		2	.40		2	-.442
7.	7	.750				7.	7	-.794
	6	.722					6	-.664

Varimax			Graphic Orthogonal			Biquartimin		
8.	13	-.724	A.	10	.71	8.	13	-.769
	16	-.556		13	.70		16	-.624
	10	-.540		11	.62		10	-.424
	9	-.511		9	.54			
	11	-.465		4	.47			
				7	.41			
				16	.41			

The first factor (C) of the graphic rotation, corresponds very closely with the first factor in both the machine rotations, but it (C) contains more highly loaded items which, in this case, facilitate interpretation. These items and their factor loadings over .40 are:

Item	Factor Loadings
7. Puts cube in cup-play	.70
8. Puts cube in cup-commission	.62
9. Scribbles in imitation	.52
6. Inhibits hand to mouth	.52
2. Can climb	.48
18. Tries to put on shoes	.42
12. Hits rubber doll	.41

This combination of behaviors seems to involve a social intercommunication type of production, a response that is culturally or socially oriented. The suggested factor name is: Cognition of behavioral units, (CBU).

The two highest loadings from all three rotation procedures coincide for the second factor with the major emphasis being upon saying four words. However, there are four items of significance in the graphic (B) rotation. These results (graphic B) follow:

Item	Factor Loading
3. Says 4 words	.91
15. Says "bye bye"	.70
6. Inhibits hand to mouth	.46
17. Bowel control	.44

Here the obvious factor meaning for the two highest loadings involves language, but the child must also express his toilet needs by verbal means. The inhibiting of hand to mouth is probably a response to parental verbal reinforcement. The language aspect in these four items is for expressing a response to a verbal communication and is here interpreted as being more semantic than behavioral in its production. The suggested name for this factor is Convergent production of semantic units, (NMU).

Factor E of the graphic rotation while containing two more high loadings seems to fit the content meaning for factor 3 on the two computer rotations. The high loadings for the graphic rotation are:

Item	Factor Loading
5. Takes third cube without dropping	.72
19. Uses spoon	.65
6. Inhibits hand to mouth	.56
11. Reproduces the form-board circle	.43

These items all involve awareness of, and relatedness to the environmental aspects of the test situation. The suggested factor name: Evaluation of figural systems, (EFS).

There is no graphically rotated factor that corresponds closely to the fourth factor of the varimax and biquartimin rotations. The biquartimin gives us only two high loadings, but the varimax (4) gives three. These varimax values are as follows:

Item	Factor Loading
4. Builds tower of two	-.799
1. Stands alone	-.496
7. Puts cube in cup in play	-.434

These items seem to involve psychomotor control although item 7 has less apparent control emphasis than do the others. Factor D of the graphic rotation which includes both Item 4 and Item 1, also includes other items that seem to have reference to psychomotor control. This graphically derived factor (D) is reproduced to enable comparison.

Item	Factor Loading
2. Can climb	.62
1. Stands alone	.61
18. Tries to put on shoes	.58
19. Uses spoon	.56
4. Builds tower of two	.45
16. Asks for things at table	.43

Factor 5 of the varimax and biquartimin contains only one high loading. This item 19 is also included in factor D of the graphic rotation. The suggested factor name for factor D which combines factors 4 and 5 is: Psychomotor control.

Factor F of the graphic rotation corresponds to the sixth factor in rotations. The highly loaded items of graphic factor F are as follows:

Item	Factor Loading
12. Hits rubber doll	.66
2. Can climb	.40

After much consideration of the possible meaning of these two items, it was concluded that the major meaning of the factor relates to memory of demonstrated past experiences. There is a behavioral aspect to these responses. A similar meaning was found in the graphic rotation, Factor E of the 12 month California Scale. The suggested factor name: Memory for behavioral relations, (MBR).

There is no graphic rotated factor that corresponds to the machine derived seventh factor. The highly loaded items for varimax were item 7, puts cube in cup, and item 6, inhibits hand to mouth. Their factor loadings were .750 and .722 respectively. In our graphic rotation, these two items belong in factor C, and as such have already been interpreted as to factor meaning.

Factor A of the graphic rotation bears similarity to the eight factor of the other rotations. Graphic Rotation A is as follows:

Item	Factor Loading
10. Unwraps cube	.71
13. Performance box	.70
11. Reproduction of form-board circle	.62
9. Scribble imitation	.54
4. Builds tower of two	.47
7. Puts cube in cup-play	.41
16. Asks for things at table	.41

These items have an element of reasoning implied in their content. The child must deduce how to carry out directions in the performance box test, and he must reason in a primitive way to know how to unwrap the cube. A suggested factor name: Deduction, Convergent production of semantic implications, (NMI).

The Merrill-Palmer Scale

Five age levels are included in the factor analyses study of the Merrill-Palmer Scale. The test records were obtained largely from the files of the Merrill-Palmer Institute, partly from the longitudinal records and partly from accumulated records from many other sources in the Detroit area. A smaller number were obtained from the files of Rachel S. Ball at Arizona State University. All age level distributions are for a two month age range and all include well over 100 test records.

Twenty-four Months Age Level. The original scale for the twenty-four months age level consisted of 24 items. Several of these were what are called all-or-none items appearing only once in the point scale. However, it was necessary to reduce the number for the final factor analysis to only 12 items because of experimental dependencies. The statistical tables are combined in Table 11 for this age level. A summary of factor loadings over .40 are given in the following tabulation matching the rotations as far as possible as to content:

Varimax			Graphic Orthogonal			Biquartimin		
1.	1	-.838	A.	1	.78	1.	2	.848
	2	-.829		2	.78		1	.839
	3	-.429						
	4	-.403						
2.	11	.643	B.	7	.48	2.	11	-.647
	10	.513		11	.47		12	-.498
	4	.484		10	.46		10	-.465
	12	.474		12	.42		4	-.442
3.	9	.842	C.	9	.91	3.	9	.887
	8	.833		8	.89		8	.843
	3	.568		4	.48		3	.522
	4	.433		3	.44			
4.	6	.714	D.	6	.68	4.	6	-.711
				3	.43			
5.	7	-.744	E.	5	.66	5.	7	.681
				7	.48			
6.	5	.710	F.	11	.46	6.	5	-.722
				4	.45			

For the first factor, both the graphic orthogonal, and the biquartimin rotations gave only two strong item loadings, while the varimax includes two others at a much lower level. The item content for the graphic rotation (Factor A) follows:

Item	Factor Loading
1. Standing on one foot	.78
2. Crossing feet	.78

In both instances the examiner demonstrates the task to be imitated by the child. Both of the items involve control of the legs since at two years of age the shortness of these appendages require an adjustment of the entire leg, and the problem really involves a kind of conscious determination of which leg to move and how to do it. The child, who is standing, has to decide how to relax one knee in order to lift one foot and stand on the other. Not much balancing is required because the response is considered a success if the child can just raise one foot momentarily. A decision must be made as to which leg is to be relaxed and lifted. A similar process is involved with crossing the feet. The child is sitting in a chair. Usually his feet stick out straight rather than reach the floor. One leg must be raised in order to place it across the other at the ankle. While there are obvious cognitive aspects to these tasks, the major implication is conscious control of the legs. Both items therefore involve proximo-distal differentiation. The suggested factor name is: Gross leg psychomotor.

There is one disagreement with respect to the second factor between the graphic rotation and the other two. Item 4, Identification of self in the mirror, is omitted, and item 7, Nest of cubes comes out with highest loading in the graphic array. While the loadings are smaller in the graphic rotation, they seem more consistent. They are as follows:

Item	Factor Loading
7. Nest of cubes	.48
11. Wallin Peg Board B	.47
10. Matching colors	.46
12. Wallin Peg Board A	.42

The related activity in these four items seems to be a kind of matching selecting the right piece to go into a particular place. The judgment required is not always the same, with the nest of cubes, it is size, the peg boards require position judgment, (one for round pegs, the other for square ones) for matching colors, it is color judgment. Our suggestion as to factor meaning is: Evaluation of figural units, (EFU).

The three rotations give identical item content for factor 3 with some difference in relative size of loading. The similar content meaning here is verbal expression. The factor loadings for the graphic rotation (c) are as follows:

Item	Factor Loading
9. Repetition of word groups	.91
8. Answering questions	.89
4. Identification of self in mirror	.48
3. Folding paper	.44

The first two items, 9 and 8, quite obviously involve language responses. The other two, with much less loading, are not so apparently language items. The mirror identification response is made vocally. The directions made for the folding paper test are demonstration along with vocalized comments. It is to be noted that both of these last two items also are included with fair sized loadings in other factors. With this in mind it is a temptation to label these items as having a verbal content and to consider verbalization as the central process involved. Actually each of the items involve behavior that Guilford would categorize as verbal comprehension. The suggested factor name: Cognition of semantic units, (CMU).

The most helpful rotation in identifying the meaning of the fourth factor is the graphic orthogonal (D). There are only two items included, each of which involves a kind of manipulation of objects.

Item	Factor Loading
6. Drawing up string	.63
3. Folding paper	.43

There are more than one type of similarity of these two items. They can involve a situation in which there must be cooperative use of two hands to achieve success. The drawing up of the string requires the alternation of two hands in grasping and pulling in the string. Folding paper requires the holding down of one edge of the paper with one hand while the other brings the opposite edge to join it, a somewhat similar process. Added to this aspect, there are also problem solving abilities that may possibly be included. Both of these items are puzzling tasks to the two year old and require effort and understanding of the problem in their solution. While there are no other even moderate sized loadings in this factor to give a cue as to which interpretation is best, Items 1 and 2 that also require coordination of the two sides of the body have the next highest loadings of .28 and .27. Hence, while not recognizing the problem solving aspect the tentative name chosen for this factor is: Motor coordination of the two hands or Hand dexterity.

Both Factors 5 and 6 for varimax and biquartimin are combined in the graphic orthogonal rotation of Factor E. The two heavily loaded items both deal with square shapes that need to be fitted together. The central adjustment is probably not judgment of size but rather the capacity to make adjustments in position by moving the wrist in order to fit the square object into a square shaped opening. The graphic rotation (E) and the fifth factor is reproduced below.

Item	Factor Loading
5. Placing 16 cubes in box	.66
7. The nest of cubes	.48

The fitting problem involves more than the mere placement of objects into a suitable container, the complication of getting all the cubes into the box, when some wrist and hand adjustment are necessary to produce the result. Hence, there is problem solving required for this 16 cubes test as well as for the nest of 4 cubes of descending size which must be fitted together. The problems are similar.

We suggest a choice of two tentative factor names: (motor aspect) Wrist and finger coordination or (problem solving aspect)

Orientation of objects in space, (CFS).

The varimax and the biquartimin both agree in giving the Factor 6 one loading over .40 and this loading is faily high, but as we have seen, the graphic rotation combined these in a meaningful way in Factor E. Factor F, however, is left with two items not much over .40 that do not seem to have any logical relationship.

Thirty Months Level. The original scale at the 30 month level consisted of 24 items. By the time all the dependent variables were eliminated however, there were only 13 left. This age level is a blending of the items more suitable for the younger ages and still usable at a higher performance level, and the beginning successes on the test items more suited to the older ages. There is therefore a wider range of test items for this level than for any of the other Merrill-Palmer age levels. The very rapid development of the young child is strikingly illustrated by the great differences in the ability between children at 24 months and at 30 months. At the older age there is much more sophistication, more comprehension of the test situation with cooperative compliance with the examiner's requests, and more genuine interest in the activities presented. For complete results of the analysis, see Table 12.

Summary of the high factor loadings follows:

Varimax			Graphic Orthogonal			Biquartimin		
1.	13	.923	B.	13	.92	1.	13	.941
	12	.524		12	.59		4	.492
	4	.524		4	.52		12	.487
	10	.484		10	.45		10	.429
				8	.42			
2.	8	-.819	C.	7	.76	2.	8	-.824
				8	.71		5	-.426
				1	.53			
				10	.41			
3.	1	-.765	A.	11	.84	3.	1	-.750
	7	-.631		9	.68		7	-.490
				5	.47			
				12	.45			
				7	.45			
				1	.44			
				6	.44			
4.	11	-.821	D.	5	.66	4.	11	.871
	9	-.705		10	.57		9	.734
	12	-.547		2	.48		12	.540
	6	-.441		3	.41		6	.438
	5	-.408						
5.	7	.665	E.	1	.46	5.	7	-.688
	10	.629					10	-.629
	5	.589					2	-.566
	3	.579					5	-.565
	2	.486					3	-.462

While not exactly in the same loading order the highly loaded items for the first factor are the same for all three rotations except for the addition in the graphic (B) of item 8, counting blocks, with a loading that is barely significant (.42). The graphically rotated (B) loadings follow:

	Item	Factor Loading
13.	Three cube pyramid	.92
12.	Little pink tower	.59
4.	cubes	.52
10.	copy a circle	.45
8.	counting blocks	.42

These items all have to do with reproducing a demonstrated production, all of them are visual in nature and involve some manipulation of test materials in the performance. The suggested name for this factor is: *Convergent production of figural systems*, (NFS).

The varimax rotation gives only one high loading, Item 8, *Counting 2 blocks* to the *second factor*, while the biquartimin adds Item 5, *Nest of cubes*. Factor C of the graphic rotation, however, gives a more complete list, if not a more meaningful one. It is difficult to assign a factor name for this rather diverse group of items. The list from the graphic rotation (C) is as follows:

Item	Factor Loading
7. Closing fist and moving thumb	.76
8. Counting two blocks	.71
1. Cutting with scissors	.53
10. Copying a circle	.41

The responses to each of these test items are of the all-or-none, non-timed variety. The first three involve demonstrated directions, but this is not done for *copying the circle*. There is production of a convergent type, and it is possibly definable as behavioral in content. The suggest name: *Production of behavioral relations*, (NBR).

The varimax and biquartimin high loadings for *Factor 3* agree with each other. Graphic rotated Factor A agrees fairly well with them, although a few more items are included. The high loadings for graphically rotated Factor A follow:

Item	Factor Loading
11. Wallin peg board A	.84
9. Seguin form board (time)	.68
5. Nest of cubes	.47
12. Little pink tower	.45
7. Closing fist and moving thumb	.45
1. Cutting with scissors	.44
6. Word groups	.44

All of these items must be comprehended by the child in terms of verbal directions as to operational procedure. Since some are timed and some are not timed, this verbal test seems to tie the meaning of the factor to cognition, rather than to speed. A suggested factor name is: Cognition of figural systems, (CFS).

Varimax and biquartimin rotations are similar in factor content of the fourth factor. Factor D of the graphic rotation corresponds fairly closely to them, and is perhaps more meaningful. Its high loading content is as follows:

Item	Factor Loading
5. Nest of cubes	.66
10. Copy circle	.57
2. One button	.48
3. Matching colors	.41

These test items all involve comprehension of the spatial position of discriminated objects. Suggested factor name: Spatial visualization, (CFT).

Factor E of the graphic rotation appears to be a residual factor, probably resulting from the fact that Factor C combined the content of Factors 2 and 3 from the varimax rotation.

The Thirty-six Months Level. Considerable difficulty was experienced at the 36 month level in eliminating the maturationally dependent items. Fifty-one original items were reduced after several factor runs to 12. The factor results are given in Table 13. The three rotations of factors for this age level follow:

Varimax			Graphic Orthogonal			Biquartimin		
1.	3	.920	D.	3	.81	1.	3	-.922
	1	.464		7	.54			
	9	.457		1	.50			
				4	.48			
				9	.46			
2.	10	.964	B.	10	.95	2.	10	-1.014
		.588		2	.51		2	- .514
				1	.50			
				12	.40			
3.	7	.666	C.	6	.62	3.	6	-.660
	4	.641		7	.57		7	-.652
	6	.611		4	.56		4	-.610
	5	.441		2	.46		5	-.435
	2	.407		5	.45			
4.	11	.619	A.	11	.59	4.	11	.628
	5	.566		2	.56		5	.472
	2	.538		5	.56		2	.446
				3	.54			
				1	.40			
5.	8	-.635	E.	8	.58	5.	8	.665
	1	-.594		1	.43		1	.533
	4	-.414						

The graphic rotation is chosen here, primarily because there are more hyperplanar values and they are more widely dispersed among the items.

One item, number 3, with very high loading in all three rotations, largely defines the meaning of the *first factor*. This item, *Closing fist and moving thumb,* describes motor coordination, the separation of thumb and fingers. The other less heavily loaded items also involve skillful use of one hand in the placement of objects involved. The graphic rotation factor D follows:

	Item	Factor Loading
3.	Closing fist and moving thumb	.81
7.	Seguin form board, errors	.54
1.	16 cubes test	.50
4.	Seguin form board, time	.48
9.	Mare and foal form board	.46

Since many children who fail the fist-and-thumb test comprehend the task but cannot wiggle the thumb without moving the fingers, the test represents finger or hand coordination rather than cognition. The other related items in this factor can very well involve the capacity to manipulate objects with thumb-finger skill, although they are sufficiently complex responses to have other factor significance. The suggested name for this factor is: Finger or hand dexterity (a form of psycho-motor ability, involving eye-hand coordination and grasping behavior.)

The three rotations agree fairly well on Factor 2, except that the graphic rotation (B) includes two other items that can be used in determination of the factor meaning. The high loadings for the graphic rotation follow:

	Item	Factor Loading
10.	Wallin peg board B	.95
2.	Repetition of word groups	.51
1.	Placing 16 cubes	.50
12.	Little pink tower	.40

The major item in this factor is Wallin peg board B, which involves the placing of 6 square pegs in 6 square holes. The nature of the response suggests that this is a temporal ordering; certainly the placing of the peg requires a time sequence. The action is sequential and requires an ordered approach. The 16 cubes also require sequential time placement of square objects. Both of these also involve wrist adjustment, but the other two items contribute nothing of this sort to the factor. The repetition of word groups would seem to involve immediate memory and the little pink tower also must be reproduced from memory. Comprehension of directions is involved but not to a differential degree for these, over other items in the factor analysis. It may be argued that the repetition of word groups involves a time sequence pattern in the repetition process, and the little pink tower also requires the placement of the five blocks of descending size, the next smaller each time in an ordering kind of activity. The very-much-to-be-questioned suggestion for a factor name is: Ordering -- of a temporal kind (NFS).

The three rotations agree with only slight differences in size of loading for the main items in Factor 3. The heavily loaded items for the graphic rotation (C) follow:

Item	Factor Loading
6. Picture puzzle No. 1	.62
7. Seguin form board, errors	.57
4. Seguin form board, time	.56
2. Repetition of word groups	.46
5. Copying a circle	.45

Picture puzzle No. 1 is a simple oval picture cut in half. The pieces of the puzzle to be put together are placed in front of the child with the instructions "put these together as nicely as you can," without implying what product is to be constructed, or suggesting that time is important in the production. Younger children who understand the instruction earlier than they get the idea of making a picture, often place the pieces in a pile, one on top of the other, hence the test is not verbally significant, since the predominating idea is that of making a whole picture out of two pieces. The speed of performance is not involved in this factor since the success of the child is included no matter how fast or slow the performance. The Seguin form board involves a rather high degree of skill and matching speed (48") that is shown by a very small percent of the three year olds. For this type of response there is a quick recognition of the form pattern of the block to be placed, and the identification of this form with the correct hole. Placement at this speed involves few errors and rather quick form matching. The copying of a circle involves reproduction of the circle pattern shown. All of these neglect the concept of speed as an essential part of the response as far as instructions to the child are concerned. Each performance involves Gestalt completion. A suggested factor name: Visual Cognition (CFU).

The two computer rotations for the fourth factor agree fairly closely with factor A of the graphic rotation, although more items are included in A, the structure of the graphic rotation is more clearly defined and simple. Hence, the Graphic factor A is presented for our consideration.

Item	Factor Loading
11. Three cube pyramid	.59
5. Copying a circle	.56
2. Repetition of word groups	.56
3. Closing fist and moving thumb	.54
1. 16 cubes	.40

4

Language repetition items in a factor with other items having form emphasis are fundamentally figural rather than semantic in nature. The child's problem is to reproduce an auditorily perceived word pattern. The three cube pyramid test involves copying a pattern of blocks from a model. The ability to reproduce the arrangement of objects in space is obviously present in the three cube pyramid. There is an ordering aspect to the performance in that the bottom blocks must be placed just the right distance apart so that the top block can be rested on them. The copying a circle is also reproduction of a sample circle placed before the child. As stated above the repetition of the word groups also involves a kind of reproduction of a vocally presented stimulus. The common element thus seems to be the reproduction of a pattern. A suggested name for this factor is: Convergent production of figural relations, (NFR).

The three rotations agree on the highest factor loadings for Factor 5. The buttoning test item has the highest loading. The graphic rotation gives the following values:

Item	Factor Loading
8. Two buttons	.58
1. 16 cubes	.43

The spatial arranging of a scrambled order is an aspect of these two items. The matching of the buttons is a task to challenge a three-year old child. One sees the child's problem best by watching him try to solve it. He often places the button-hole strip very carefully on the button strip, matching the button with the hole. When this method is used, success is much more likely than when he buttons the first one indiscriminately with whichever button he chances to contact with a button-hole. Success with the 16 cubes also involves placing in an orderly fashion the blocks in their box. Hence the factor name may be: Cognition of figural systems, or spatial orientation, (CFS).

Merrill-Palmer, Forty-eight Months Level. At this age level there were originally 59 items. They were reduced, because of experimental dependencies, to 16 items in our final factor analysis. There is some disagreement between the graphic, and the other two rotations for this 48-month level study. Following is a list of rotated factors in which the divergence in factor content among the rotations can be seen.

Varimax			Graphic Orthogonal			Biquartimin		
1.	10	.778	A.	16	.64	1.	10	.765
	6	.724		10	.61		16	.746
	9	.714		11	.60		9	.727
	11	.618		9	.59		11	.634
	1	.552		14	.42		13	.561
	14	.527					1	.481
	13	.459					14	.469
2.	5	-.858	B.	5	.64	2.	12	-.855
	12	-.822		4	.63		5	-.833
				2	.50			
				3	.42			
				7	.41			
3.	3	.719	C.	12	.90	3.	7	-.721
	7	.714		5	.58		3	-.701
				13	.53			
4.	8	.761		15	.48	4.	8	.777
	15	.599		9	.40		15	.556

Varimax			Graphic Orthogonal			Biquartimin		
5.	6	.559	D.	8	.70	5.	6	.538
				15	.63			
				4	.48			
				13	.43			
6.	4	-.740	E.	3	.71	6.	4	-.758
	2	-.545		1	.51		2	-.531
	1	-.525		10	.45		1	-.490
	14	-.477					14	-.448

In the graphic rotation only five principal component factors came through. The sixth seemed to have no potential for differentiating meaningful clusters. Its loadings were below .35 in absolute value. After quite a bit of exploration, the number of items in hyperplanes seemed to be about five, so to use more factors would make the solution more indeterminate.

The graphic rotation (A) seemed to be more meaningful for the first factor. The high loadings are as follows.

Item	Factor Loadings
16. Action Agent	.64
10. Decroly matching	.61
11. Manikin	.60
9. Picture puzzle No. 3	.59
14. Mare and foal	.42

All of these items involve knowing the usual relations within a system, what goes with what. The Action agent has been linked consistently with the Decroly matching test in every factor analysis made with the Merrill-Palmer Scale. The inclusion of the Manikin which involves putting parts of a man together and the Picture puzzle No. 3, which is a fairly difficult four piece jig saw puzzle, not in a frame, strengthen the conclusion that this factor does not emphasize the verbal aspects of the Action agent, but rather involves seeing relationships among parts and the working out of what parts belong together. The matching and form completion is in terms of the meaning of the forms, perhaps even verbalized and the Action agent questions involve this same process, matching the question with the answer, (e.g. Action agent questions, What sleeps? What swims?). The inclusion of the Action agent with a high loading would stand against the probability that motor skills or spatial judgments were important in the factor meaning.

The suggested factor name is: Cognition of semantic systems or semantic relations, (CMS or CMR).

Graphically rotated factor B includes five items whereas the varimax on the biquartimin procedures each allowed only two items for the second factor. There is therefore, a better chance of clarifying the meaning of the factor as being something more definitive than simply the ability to put picture puzzles together which is what is indicated by the varimax and biquartimin rotations. The items included in the graphically rotated (B) factor are as follows:

Item	Factor Loading
5. Picture puzzle no. 1	.64
4. Copying a cross	.63
2. Opposition of thumb and fingers	.50
3. Four buttons	.42
7. Buttoning two buttons	.41

The thumb and finger opposition and the buttoning tasks require coordination of small muscles, but the two heavier loaded items seem to indicate that the major process implied by this factor is cognition involving spatial orientation or relations.

Certainly opposition of the thumb and fingers also requires the child's realization that the fingers must touch the thumb in a sequential rotated order (haphazard responses are scored as failures, and, as suggested in our earlier discussion of the two button test, spatial orientation is involved as well as coordination of the two hands. The suggested factor name is: Spatial orientation, (CFS).

Again in factor 3 the graphic rotation gave a larger number of items with more clues for a determination of the factor meaning than do either of the other rotations. The large factor loadings for graphic rotation C are as follows:

Item	Factor Loading
12. Picture puzzle no. 2	.90
5. Picture puzzle no. 1	.58
13. Seguin form board (time)	.53
15. Six-cube pyramid	.48
9. Picture puzzle no. 3	.40

The presence of all three picture puzzles in this factor might indicate the possible compounding of a "format" factor as discussed by Cattell (1958), and a pattern completion aptitude. However, the inclusion of Items 13 and 15, Seguin form board and six-cube pyramid, requires more than a "format" interpretation. These two items also have another factor in common (D) which is not shared by the picture puzzle tasks. Suggested factor name: Convergent production of figural relations, (NFR).

The two high factor loadings in the computer rotations for factor 4 are also the highest loadings in the graphic rotation, figure D, but in the graphic rotation two other items are included to give more definitive meaning. The leading tests involve comprehension of three-dimensional relationships plus an immediate memory of the form pattern essential for the pattern completion. The suggested factor name: Memory for figural systems, (MFS).

A factor 5 containing only one high loading, Item 6, the three-cube pyramid, is provided by both of the computer rotations. The graphic rotations, however, places Item 6 in factor D with a loading of only .37. But the fifth factor in the graphic rotation, Factor E, offers more potential meaning. The high loadings are:

Item	Factor Loading
3. Four buttons	.71
1. Seguin errors	.51
10. Decroly	.45

The leading test item (the Four Buttons Test) has been, in other instances, interpreted as "ordering," but what all three items of this factor seem to share is a kind of foresight in the sense that if the child makes a mistake early, his "degrees of freedom" are reduced. No portion can be used twice. The suggested factor name, therefore, is Perceptual foresight, (CFI).

It is interesting that the Seguin time and error scores have different factor locations. The original scale included both time and errors because they were found to have rather low correlations with each other. There are comparable findings in other studies showing analogous lack of correlation between speed and comprehension.

The Fifty-four Months Level. Five factors were extracted for this level of test items (Table 15). There was unusual agreement in factor content for the three rotations and they were fairly simple in structure. However, five of the twelve items were complex and appeared in at least two of the factors. A list of the items having factor loadings of over .40 for the three rotations follows:

Varimax			Graphic Orthogonal			Biquartimin		
1.	9	.600	D.	12	.62	1.	9	-.493
	12	.527		9	.58		12	-.460
				4	.52			
				1	.52			
				7	.41			
2.	1	.912	C.	2	.78	2.	1	-.945
	2	.844		6	.52		2	-.873
				1	.46			
3.	4	-.849	A.	8	.77	3.	4	-.916
	8	-.713		10	.56		8	-.751
	10	-.552		4	.54		10	-.565
	5	-.516		5	.51		5	-.477
	7	-.438					7	-.436
4.	11	-.842	B.	11	.76	4.	11	.904
	5	-.636		5	.56		5	.595
	9	-.498		9	.50		9	.453
	6	-.445		6	.46		6	.416
5.	3	-.625	E.	3	.62	5.	3	-.565
	8	-.413		1	.46			
	2	-.406						

The most heavily loaded items for the first factor of the two computer rotations agree with the graphically rotated factor D. Since factor D has other lesser loadings that assist in determining the factor definition, it is listed below for our consideration.

Item	Factor Loading
12. Manikin	.62
9. Seguin form board	.58
4. Picture puzzle #2	.52
1. Four buttons	.52
7. Copy star	.41

All of the items have to do with spatial judgments in which some operation is performed that produces a modification of the test materials in an altered space relationship. The suggested factor name: Spatial orientation, (CFT).

The varimax and biquartimin rotations give to the second factor only two items, each with very high loadings. The corresponding graphic rotation (C) includes the 6 cube pyramid test item. Graphic factor (C) is presented here:

Item	Factor Loading
2. 3 cube pyramid	.78
6. 6 cube pyramid	.52
1. 4 buttons	.46

The one-inch cubes in the pyramid tests are built and left standing for the child to copy with other cubes. The button test involves placing the button and button-hole strips together so that they can be buttoned in the correct sequence. The response seems to require the cognition of Gestalt patterns. Suggested name: Cognition of figural relations, (CFR).

Graphically rotated factor A agrees fairly consistently with the two computer rotations of factor 3, except that Picture puzzle no.2 is de-emphasized. Graphic factor A is here interpreted as the third factor.

Item	Factor Loading
8. Action agent	.77
10. Decroly matching	.56
4. Picture puzzle no. 2	.54
5. Mare and foal	.51

The combination of items for this factor is very similar to factor 1 at the 48 month level, and the discussion of the Action agent and the Decroly matching game is also pertinent to this very similar factor. The difference of the Picture puzzle no. 2 instead of no. 3 and the elimination of the Manikin item are relatively unimportant. The Mare and foal test is also included in both factors. Suggested factor name: Cognition of semantic systems, (CMS).

The items included in factor 4 with loadings over .40 are exactly the same, and in the same order for both varimax and biquartimin rotations, and for factor B in the graphic rotation. The graphic factor B values are given:

Item	Factor Loading
11. Little pink tower	.76
5. Mare and foal	.56
9. Seguin form board	.50
6. Six-cube pyramid	.46

These items all have the common quality of requiring a pattern completion aptitude. The suggested factor name: Figural redefinition, or convergent production of figural systems, (NFS).

The varimax rotation gives more content to factor 5 than do the other two. The test item with the highest loading, however, is the same in all three rotations. The graphic rotation

(E) gives easily interpretable results, hence the graphic rotation follows:

Item	Factor Loading
2. 2 buttons	.62
1. 4 buttons	.46

This is possibly a task-specific factor involving the skill of buttoning, and not a generalized ability. These two items are partly dependent variables, and Cattell would probably identify them as producing a format factor. The probability of passing the two button test at the difficulty level (item 75) of 23 seconds or less, when the four button test (item 55) at the 76 second performance level has been failed, is .237. This would indicate that there is only partial dependency of the items. Since the task does involve the coordination of two hands, the suggested factor name is: Finger or hand dexterity.

The Stanford-Binet (Form L) Scale

Three age levels were selected for study of the preliterate portions of the Stanford-Binet (Form L) Scale. In the initial stages of our study the item lists were all heavily laden with dependent variables that had to be eliminated. Dropping these items necessarily reduced the range of difficulty covered by the test. We were forced to reduce the three-year level from an original list of 24 variables to 15 for the present factor analysis. The four year level had 30 variables to start with and ended up with only 12 variables. The five year level also had 30 variables to start with and all but nine were omitted. This elimination of dependencies should make a clearer factor picture of this test scale than has heretofore been presented.

The number of cases for the two older age levels was below our standard, because of the limited number of test protocols available to us at these two narrow age levels, either in our Merrill-Palmer files, or in those of the Arizona State University Psychological Clinic.

While the number of dependently related tests that we omitted was large, even among the test items retained at certain age levels, the tests fall largely into one factor, making an obviously narrow coverage of ability areas. There is no indication that the difficulty level of the test items, as shown by the computed probabilities of failure, enters into the factor picture.

The Three Year Age level. The three year level of the Stanford-Binet (L) shows considerable variety and many less experimental dependencies than at the other two age levels of the scale (Table 16). A comparative list of the factor loadings follows:

Varimax			Graphic Orthogonal			Biquartimin		
1.	15	-.885	A.	15	.82	1.	15	.889
	13	-.498		13	.49		4	-.529
	9	-.470		4	-.46		13	.478
	4	.464		9	.45		9	.434
2.	2	.877	B.	2	.97	2.	2	-.880
	7	.667		8	.68		7	-.840
				7	.60		12	-.536
				9	.55			
				13	.47			

Varimax			Graphic Orthogonal			Biquartimin		
3.	1	-.732	C.	3	.73	3.	1	-.921
	3	-.696		1	.47		5	-.779
	5	-.692		5	.46		3	-.611
	4	-.413						
4.	11	-.855	D.	11	.91	4.	11	.919
	14	-.691		14	.73		10	.655
	10	-.617		10	.55		14	.628
	12	-.550		12	.53		2	-.403
	5	-.460		3	.47			
	3	-.449		13	.41			
				9	.40			
5.	6	-.747	E.	6	.79	5.	6	.582
				5	.64		3	-.531
				1	.59			
				10	.41			
6.	8	.687				6.	8	.713
	13	.557					13	.582
	9	.473						
	11	.423						
	2	.408						

For the first factor, the four high factor loadings for the three rotations agree except for slight reversals in order of size. The graphic rotation order and loadings are given here:

Item	Factor Loading
15. Three commissions	.82
13. Comprehension II	.49
4. Obey simple commands	-.46
9. Comprehension I	.45

It is noteworth that the item, Obeying simple commands, placed at Year III-6 has a negative factor loading in contrast to the other significant items in this factor. What are the differences that may account for this negative loading? The directions for this test item allow for repeated trials. The authors state, "It involves comprehending directions. The tasks are so simple that recall is probably not a factor in success at this level" (Terman and Merrill, 1937, p. 203). The item is presumed to be a simplified form of the Three Commissions occuring at Year IV-6. The simpler task involves performance of only one commission at a time, and the object to be manipulated is selected from a row of five objects. The Three Commissions test, on the other hand, requires no discrimination of objects, since they have been pointed out and named in the directions, but rather it is a memory ordering task. "Success in it depends chiefly on the ability to keep the three commands in mind long enough to carry them out without confusion or errors" (ibid., p. 215). Success requires that all three commissions are executed in the proper order, no prompting or repetition of directions is permitted. Thus, it can be seen that the later task, in spite of the author's intentions, is not a more difficult version of the earlier one.

The Comprehension Tests I and II at two difficulty levels require the child to "comprehend the situation suggested" (ibid., p. 206). If the child singles out some detail and ignores completely the place of the part with reference to the whole, he fails the test.

Thus the similarity between the Three Commissions and the Comprehension tests is somewhat obvious. There is a need to keep in mind the whole idea. Response to a segment of the task produces failure. Furthermore, the task must be comprehended with one presentation. Immediate memory is therefore also involved.

In the case of Obeying Simple Commands successful performance involves the selection of a part of the whole; immediate memory is not especially involved. The task in this case is therefore opposite to that required for the three other items in this factor.

The factor content can then be interpreted to involve memory for a verbal communication involving a decision to make, what to the child is an appropriate response. The suggested factor name is: Memory for a Semantic system, (MMS).

The five items of the graphic rotation for factor 2 are the most helpful in determining a possible factor meaning. These items are as follows:

Item	Factor Loading
2. Copy circle	.97
8. Identify object by use	.68
7. Response to pictures	.60
9. Comprehension I	.55
13. Comprehension II	.47

Each of these items involves Gestalt completion, as presented before in similar factors, the name given by Guilford to this process is Visual Cognition, (CFU).

The varimax rotation gives one more loading over .40 than the other two rotations, but otherwise the high loadings for factor 3 are the same for each type of rotation. The graphic rotation content is given here:

Item	Factor Loading
3. Repeats 3 numbers	.73
1. Picture memories	.47
5. Picture vocabulary	.46

These items all involve memory. The Picture Vocabulary, with the lowest significant loading, requires the child to identify objects by naming them, a memory task. The suggested factor name is: Memory for symbolic systems, (MSS).

The high loading content for the graphically rotated factor (D) we present here as representing the fourth factor

Item	Factor Loading
11. Picture identification	.91
14. Picture likenesses and differences	.73
10. Naming objects from memory	.55
12. Discrimination of forms	.53
3. Repetition of three numbers	.47
13. Comprehension II	.41
9. Comprehension I	.40

The visual aspect of these tasks is indicated in all but the last three of the items. All of them, however, involve naming responses. A suggested name for the factor is: Concept naming, (NMU).

The last two factors in the three factor rotations are a bit confused. The graphic rotation accounts for the variance in five factors, the other two have six factors.

The fifth factor in the varimax array has only one significant loading, while in the biquartimin it has two. The graphic array, on the other hand, presents four significant loadings, with more easily interpretable meaning. The graphic rotation (E) is reproduced:

Item	Factor Loading
6. Comparison of sticks	.79
5. Picture vocabulary	.64
1. Picture memory	.59
10. Naming objects from memory	.41

These items all involve evaluation and this may be the implication of their combination in this factor. A suggested factor name is: Evaluation of figural systems, (EFS).

Although the two computer rotations came out with significantly high loadings for a sixth factor, the graphic procedure found only a residual factor F with no high loadings (Table 16).

The Forty-eight Month Level. The correlation matrix for the preceding 36 months age level gave a wide range of values and the factor results showed much variation in content. The other two levels of the Stanford-Binet (L) (48 and 60 months) were much less diversified. In the analysis of the 48 month level (Table 17), a number of items were so experimentally dependent that they could not be used for factor analyses. Only 12 items could be retained out of the original 30. Out of 66 intercorrelations there were only 8 that were less than .30, while 31 were higher than .50. Only 4 factors were extracted. In this analysis the graphic rotation varied more from the varimax and biquartimin rotations than in most of the other age levels studied. Because of the high relationship between items there was more tendency toward general factors.

A re-examination of the analysis of the Cattell tests (Tables 5 and 6) will show that they are more similar in structure and in narrow meaning range, to these two older Stanford-Binet age levels, than are any of the other age levels, including the much more adequate Stanford-Binet (1) age level of 36 months. The factor contents of the three rotations are sketched below, matching factor content where possible.

Varimax			Graphic Orthogonal			Biquartimin		
1.	11	-.836	C.	9	.75	1.	11	-.971
	5	-.736		11	.74		5	-.732
	12	-.569		12	.50		3	-.573
	3	-.548		5	.41		12	-.558
	6	-.539					6	-.502
2.	2	-.973		2	.76	2.	2	-1.032
	4	-.589		3	.50		4	- .410
	3	-.503						
3.	10	.773	A.	10	.84	3.	8	.859
	8	.760		5	.80		1	.818
	1	.751		1	.74		10	.809
	4	.534		6	.70		4	.557
	5	.491		8	.69			
	6	.429		4	.62			
	7	.413		3	.60			
				12	.58			
				7	.56			
				11	.46			
4.	9	-.767	B.	2	.57	4.	9	-.644
	11	-.404		8	.45			

Factor 1, as usual, has similar content in the two computer rotations. There is some similarity of the corresponding graphic rotated factor C with the other two. Graphic factor C, representing the first factor, is presented here for our study.

Item	Factor Loading
9. Three commissions	.75
11. Paper folding	.74
12. Counting 4 objects	.50
5. Aesthetic comparison	.41

All of these items seem to involve selecting the best solution to a practical problem or practical judgment. The factor name suggested, therefore, is Evaluation-practical judgment, (EMT).

The varimax rotation includes in its second factor the items covered in the other two rotations most nearly like it, namely factor D and biquartimin factor 2. For our interpretation of this factor we present the highly loaded items of factor D of the graphic rotation.

Item	Factor Loading
2. Naming objects from memory	.76
3. Picture completion (man)	.50

The basic meaning is indicated by the high factor loading of Item 2, naming objects from memory. The other item requires meaningful memory also, but is more complex in its factor meaning than Item 2. The suggested factor name: Meaningful memory, (MMR).

All of the items with high loadings in both of the machine rotations for the third factor, also have high loadings in graphic factor A. Graphic A is reproduced here with its long list of items with high factor loadings, which include all but two of the variables.

Item	Factor Loading
10. Opposite analogies	.84
5. Aesthetic comparison	.80
1. Comparison of sticks	.74
6. Repeating 4 digits	.70
8. Materials test	.69
4. Comprehension II	.62
3. Picture completion (man)	.60
12. Counting 4 objects	.58
7. Picture likenesses and differences	.56
11. Paper folding	.46

All of these items require the process of comparison on the part of the child, especially the three with highest loadings. Suggested factor name: Cognition of semantic relations, (CMR).

The graphic rotation seems to have absorbed into its Factor C, items 9 and 11, which appear in factor 4 of the varimax rotation, while splitting item 2 between its two factors B and D. Because of the confusion still present it seems that the most meaningful rotation is the Graphic B for this fourth factor.

Graphic factor B is given below:

Item	Factor Loading
2. Naming objects from memory	.57
8. Materials test	.45

Both heavily loaded items for this factor involve memory for temporal order and are given the tentative factor name: Memory for temporal order, (MMS).

The Sixty Months Level. Contrary to our experience with many of the other age level analyses, the graphic rotation was least satisfactory for all the factors at this age level in clarifying possible factor meaning. However, for the sake of consistency the Graphic rotations are used for determining the factor names. The complete results of this analysis are given in Table 18. The summary of factor loadings over 40 for each of the rotations follows:

Varimax			Graphic Orthogonal			Biquartimin		
1.	2	.747	B.	8	.84	1.	2	-.830
	3	.591		7	.83		3	-.547
	8	.562		3	.80		8	-.490
	9	.460		9	.78		9	-.403
	5	.422		4	.78			
				5	.71			
				6	.58			
				2	.57			
				1	.40			
2.	1	.906	A.	1	.77	2.	1	.903
	4	.737		4	.60		4	.409
	5	.567		5	.40			
3.	7	-.718	C.	2	.54	3.	4	-.832
	4	-.663					7	-.579
	9	-.569					9	-.422
	8	-.524						
	3	-.480						
	5	-.425						
4.	6	.612	D.	6	.56	4.	6	.679

As can be seen from the correlation table (Table 18) for this age group, the correlations are very high with only eight below .50. The assumption has been that if all the test items correlate highly, thus producing a general factor, that this factor is "general intelligence" (g).

However, the presence of the general factor merely means that the range of meaning included in the list of test items is extremely narrow. It does not necessarily imply a broad general meaning.

Since factor B of the graphic rotation presents all the items as having loadings over .40, these can be called a general factor, but such a designation hardly yields adequate understanding in regard to the factor content of the first factor. A more meaningful interpretation is that the breadth of mental abilities covered is much restricted. By utilizing the items having high loadings in the biquartimin and the varimax as cues in regard to the content meaning, we find that all the items can be pinned down more definitely to a Gestalt completion, or speed of closure implication. The structure of intellect term for this is: Visual Cognition, (CFU).

There is agreement in the high loadings of factor 2 of the varimax with factor A of the graphic rotation. The item loading values for graphic factor A follow:

Item	Factor Loading
1. Three commissions	.77
4. Memory for sentences	.60
5. Vocabulary	.40

These items seem to refer to some common memory aspect involving memory for ideas verbally communicated. The suggested second factor name, therefore, is: Memory for ideas, (MMU).

The graphic rotation gives only one item for factor C, the third factor (item 2, picture completion, factor loading .54). This item has the highest loading in Factor 1 of both the varimax and the biquartimin rotations. These two computer rotations are in agreement for the highest four loadings in factor 1. However, since the graphic rotation is used in interpreting the other factors in this factorization, (factor C) is considered here. The rank order of the loading values in graphic C, although all are under .40, are in agreement with the order of values for factor 1 in the two computer rotations. With only one high loading for factor C, it is unsatisfactory to assign a factor name. However, the similarity in relative values of the items to the other two rotations, gives cues for the possible meaning of this factor. All of the high values for Varimax and Biquartimin factor 1 are concerned with spatial judgments, and certainly this is one important facet of the picture completion item. Therefore the tentative factor name suggested is Spatial visualization, (CFT).

For the fourth factor, all three rotations give only one high factor loading for the same item 6, copying a bead chain from memory. With only one high item loading, one does not know what meaning to ascribe to the factor, since there are several potential interpretations to this task.

A guess as to the meaning is: Memory for figural units, (MFU).

Discussion of Factor Findings

In the preceding sections we have attempted careful analyses of the various functioning processes that make up the item content of infant and preschool tests at the different age levels. The tentatively assigned names have been suggested separately by various individuals who have analyzed the potential meaning, and then in conference with each other, the naming process was continued until there was agreement on the factor names. The final names suggested are logical to persons who know the test content. They are not presented here as the only valid names that can be given. Further studies are needed to determine more definite factor meanings. When a sufficient number of factorizations have been made at preschool and infant levels the identification of factors can be made as meaningfully as they are for older persons. Of course, a certain amount of "guess-work" is necessary in the naming of the factors in preschool and infant tests at the present time.

Because of the natural uninhibited responses of the young child, mental test results that pool all the varied responses into one unitary score such as an IQ, give an inadequate picture of the child's performance. Certainly a profile of ability based upon the factor significance of the items is readily obtained and offers far more meaning for the mental test evaluation. The long felt need for clinical testing tools giving part score meaning has produced devices for part scores that are based upon theoretical concepts such as performance ability versus verbal ability. Actually, as can be seen by the present study, the division of tests into verbal and non-verbal items are not necessarily discriminative. Verbal tests may involve the same processes as the so-called motor tests. Tests requiring responses using words may really emphasize memory, or evaluative abilities that cannot be classed as semantic. Tests dealing with verbal content, without any objective test materials and those involving manipulation of blocks or form fitting situations, may both call for the same kind of cognitive response. One cannot by inspection or observation alone discover the meaning of items. This point is particularly shown by the Gesell developmental items where inspection of the factor results for one factor give a wide variation in the original Gesell classification of items.

Item Name and Symbol Classification		Factor Loading
P 41 (A)	Puts cube in cup-play	.70
A 40 (C)	Puts cube in cup-commission	.62
M 40 (B)	Scribbles in imitation	.52
M 34 (B_1)	Inhibits hand to mouth	.52
M 20 (B+)	Can climb	.48
P 20	Tries to put on shoes	.42
A 22 (A)	Hits rubber doll	.41

Here the responses seem to be culturally or socially oriented. The suggested factor name is Cognition of behavioral units, and yet there are some items from all four of the Gesell categories. Three of these with fairly high loadings have been labelled "motor" (M). All six factors of the graphic rotation for this matrix contain at least one of these M items. Therefore, to pool all of these items, and from them to evaluate only the child's motor development, is to misinterpret the meaning actually involved in the items.

Yet the standardized test situation can afford a quick and adequate means for the evaluation of the child. The contribution to the understanding of the child by observing his varying responses in a controlled environment by an experienced and insightful examiner, cannot be denied, but without a standardized test utilizing all the potential assistance possible in making the evaluation, the results are largely opinion and subjective judgment. Such findings are of value only in so far as the examiner is capable of using observational cues. They have little to commend them in terms of objectivity. The results may be meaningful and valuable to an experienced clinician, but without standardization there is failure in communication with others who need to apply the test findings. However, this kind of evaluation is to be preferred to that of the examiner who simply follows the rules laid down by the test directions and comes out with a generalized score, serenely sure that the measure is an objective one.

If mental tests can be devised, utilizing the presently available knowledge about test development such as item selection, factorial validities, modern scoring devices that are statistically sound, and up-to-date standardization procedures, we shall have made a new step forward in providing more truly objective and trustworthy tools for clinical use.

In the evaluation in our present study of the operations elicited by the test items, we discover cognition, memory, evaluation, and both divergent and convergent production in addition to the hand dexterity and other psycho-motor skills for which we found no ready intellectual name. It is felt that the resort to the naming of test performance as "psycho-motor" is an acknowledgement of failure to analyze the behavior adequately, and thus emphasizing the motor, rather than the *psychological* implications of the task. Most of the reactions of the child require intellectual processes to accompany the movements involved. They are goal-directed, examiner directed, and involve discrimination, selection and control of the neuro-muscular responses to be used. They are more appropriately chosen actions for the age level studied than they would be for older ages where the behaviors have become more habitually performed. These tests, involving what seem like simple motor responses at these young ages, are actually much more complicated and more needful of selective judgment and control than we realize. This is why they are indicators of intellectual ability. The habitual responses of children at more mature levels have not been developed at these younger ages. Superficial judgment about these responses, lumping them into a category called *psycho-motor* or *performance*, constitute a very inadequate analysis of the behavior involved.

The factor names for each of our fourteen age level scales have been summarized below. The factor number is given, along with the proposed factor name, using the Guilford Structure of Intellect classification label where possible.

Cattell 3

1.	Attention of a visual type, Visual Cognition	CFU
2.	Convergent production of a hand-arm type to a visual stimulus	NFR
3.	Divergent production with fingers, Figural Fluency	DFU
4.	Perceptual foresight - Cognition	CFI

Cattell 6

1.	Cognitive - unilateral reaching for object	CFU
2.	Experiential evaluation	EMS
3.	Convergent production of figural units	NFU
4.	Divergent production - exploratory manipulation	DFT

California 6

1.	Visual cognition of figural relations	CFU
2.	Cognition of behavioral units	CBU
3.	Evaluation of semantic systems	EMS
4.	Divergent production of figural relations	DFR
5.	Evaluation of semantic relations	EMR

California 12

1.	Deduction - Convergent production of semantic implications	NMI
2.	Convergent production of figural units	NFU
3.	Cognition of behavior units	CBU
4.	Cognition of figural relations	CFR
5.	Memory for figural relations	MFR
6.	Memory for behavioral relations	MBR

Gesell 6

1.	Divergent production of figural relations	DFR
2.	Meaningful memory	MMR
3.	Experiential evaluation	EMS
4.	Locomotor ability	Locomotor
5.	Reflex development	Reflex
6.	Divergent production of figural units	DFU
7.	Convergent production of figural relations	NFR
8.	Whole-body control	Whole-body

Gesell 12

1.	Cognition of behavioral units	CBU
2.	Convergent production of semantic units	NMU
3.	Evaluation of figural systems	EFS
4&5.	Gross psycho-motor control or residual gross psychomotor	Gross psycho-motor
6.	Memory for behavioral relations	MBR
7.	Deduction: Convergent production of semantic implications	NMI

Merrill-Palmer 24

1.	Gross leg psychomotor	Gross psycho-motor
2.	Evaluation of figural units	EFU
3.	Cognition of semantic units	CMU
4.	Hand dexterity	Hand dexterity
5.	Spatial orientation	CFS
6.	Residual	

Merrill-Palmer 30

1.	Convergent production of figural systems	NFS
2.	Convergent production of behavioral relations	NBR
3.	Cognition of figural systems	CFS
4.	Spatial visualization	CFT
5.	Residual	

Merrill-Palmer 36

1.	Finger or hand dexterity	Hand dexterity
2.	Ordering of a temporal kind	NFS
3.	Visual Cognition	CFU
4.	Convergent production of figural relations	NFR
5.	Cognition - spatial orientation	CFR

Merrill-Palmer 48

1.	Cognition of semantic systems	CMR
2.	Spatial orientation - cognition	CFS
3.	Convergent production of figural relations	NFR
4.	Memory for figural systems	MFS
5.	Perceptual foresight	CFI

Merrill-Palmer 54

1.	Spatial orientation	CFT
2.	Cognition of spatial relations	CSR
3.	Cognition of semantic systems	CMS
4.	Convergent Production of figural systems	NFS
5.	Finger or hand dexterity	Hand dexterity

Stanford-Binet (L) 36

1.	Memory for a semantic system	MMS
2.	Visual Cognition	CFU
3.	Memory for symbolic systems	MSS
4.	Concept naming	NMU
5.	Evaluation of figural system	EFS
6.	Residual	

Stanford-Binet (L) 48

1.	Evaluation - Practical judgment	EMT
2.	Meaningful memory	MMR
3.	Cognition of semantic relations	CMR
4.	Memory for temporal order	MMS

Stanford-Binet (L) 60

1.	Visual cognition	CFU
2.	Memory for ideas	MMU
3.	Spatial Visualization	CFT
4.	Memory for figural units	MFU

While there are many duplications in factor names, the age levels show much diversification in the names assigned. For comparison of the factor meaning at the different age levels, the Table 19 is given.*

Comparison of Factor Meaning. Let us analyze this table to see what similarities and differences exist in factor content. Out of 72 factors, 33 percent are cognitive, 14 percent are memory,

---

*We are indebted to Phillip R. Merrifield for this table which has been adapted to our use in this discussion.

8 percent are divergent production, 19 percent are convergent production, 10 percent are evaluation and 16 percent are classified as involving either some form of psychomotor response or are residual. This means that since over a third of the factors are cognitive in nature the other aspects of intellect are somewhat slighted. Only 13 percent of the factors involve memory. For the infant scales only the 12 month California and the Gesell schedules include a memory factor. The Merrill-Palmer Scale has little memory content while the Stanford-Binet has two memory factors for each age level. It is enlightening to see that while the infant scales have considerable reference to divergent production, the operation involving creative abilities, the preschool scales have no factor involving this type of response. Evaluation is also rather inadequately represented. There is only one factor in the Merrill-Palmer Scale that isolates this operation. Hand dexterity is well represented by the Merrill-Palmer Scale, but it is probably possible to label these items by other more meaningful names than this rather generalized, but essentially facilitating, trait.

It is clear from Table 19 why different tests of preschool intelligence fail to correlate very much with each other or with tests of later ages that emphasize verbal (semantic) tests. Of these younger test levels, only the Binet has anything but token representation in semantic items, and it lacks almost entirely figural content. The Gesell levels seem closest to the Binet in a gross approximation. The Merrill-Palmer levels are predominantly figural and symbolic, almost the converse of the Binet.

The Merrill-Palmer Test also has the best showing in convergent production although it is at times difficult to differentiate convergent production from evaluation, which is scarecely included in The Merrill-Palmer Scale.

Several haunting questions arise as we study this table of comparative factor content. Why should CFS show up at four Merrill-Palmer age levels and not at 54 months? Why should it have only one memory factor at just the 48 months level, and only one evaluative item, EFU, at the 24 months level? The conjecture is that the representative test items selected at these different age levels differ in content. Surely the 24 months level has many different items from the age levels over three. Also the items at 30 and 36 months differ from the older age levels, although some appear consecutively at three or more levels. The 48 months level includes five items not in the 54 month level and nine items not in the 36 months level. There are only five items at the 24 months level that appear also in the 30 months level. Hence one explanation can be that the content is different at the various levels. However, the content value of the questions also change as the child grows older and they may not test the same processes at different age levels.

There is no criterion as to what the factor structure should be in terms of the various operations. The proportion of cognitive factors should very well be greater than the others. It seems that all these varying operations require also cognitive elements. For example, meaningful memory, perceptual speed, concept naming, and logical reasoning all involve cognition, but the major process selected by the factor in instances where these are preponderant can be identified as having more prominently these other intellectual operations. Similarly, it can be stated that all the necessary orientation for performing any intellectual operation involves a background of meaningful memory, but the major process involved in the factor content may be other operations than the contributing memory aspect.

We can see that the present mental test scales for the infant and preschool levels deal preponderantly with the cognitive and convergent production aspects of intellect. If this is true of older age levels, then it is small wonder that the IQ has little relationship to creativity which is best represented by the divergent production factors.

The differences between the infant tests and the later preschool scales in types of operations indicated by our factor names is also striking. The infant scales are more diversified and cover a wider range of intellectual ability. The preschool tests are similar in content except that the Merrill-Palmer includes more reference to manual skill and the Stanford-Binet has more factors involving memory. The Stanford-Binet at 60 months particularly is narrow in its meaning span, with two cognitive factors and two memory factors, only one of which is semantic. Three are figural and one is semantic in content and it has no evaluative or production operations.

Since many of the test items were eliminated because of the dependent relationship to other items, it might be inferred that we have by this procedure reduced the breadth of the factor structure of the various age levels. However, in each instance one of the dependent variables was retained, so that duplication of test item meaning was reduced, and more simple factor structure was gained by this process. The actual factor significance presumably has been retained and the naming process has been simplified.

Considering these differences in test structure, the variations in score from one test scale to another is to be expected. This does not necessarily mean that one test has no value because it does not correspond in IQ with another test taken at the same time. It is almost surprising that there is a degree of correlation between two simultaneously administered scales as seems to be present. When considerable time elapses between the test and retest, even two levels of the same scale have somewhat different factor structure. This then is one of the reasons why the retest picture shows so much variability. Of course, the differences in test environment, the child-examiner relationships, and the child's own behavioral variability are also partly responsible for the striking variations found in retests.

It must be borne in mind in making a study of intellectual development of children through the use of mental tests, that the results obtained are determined by the kinds of tests used and their adequacy in discovering the potential range of abilities. There are many other possible insights about children in test situations than are shown in this study of the success or failure of separate test entities, and undoubtedly concealed in these data are truths which because of the nature of our records, lie buried. Such observations as the amount of interest and motivation called forth by the test, the degree of dependency of the child upon the examiner, the carefulness, the orderliness, the spontaneity of the response, and the originality shown, are all possible kinds of test record information.

The pooled data without cues separating the varied behaviors of different children give no picture of these qualities. In our report it is possible to study the item response in terms of intellectual operations only because the cues recorded were in regard to such behavior.

Conclusions based upon tests covering a restricted variety of items with high inter-correlations will seem to show a general factor, while those having wider coverage will reflect this breadth of content. Tests emphasizing verbal elements will seem to indicate preponderance of this type of development. Those emphasizing cognitive or productive aspects will tend to make the child seem to have a preponderance of these abilities. In each case it will be the items involved and the child's response to them that is really recorded and not a direct evaluation of the intellectual capacity of the child himself.

VI

SUMMARY AND CONCLUSIONS

Ever since Binet began his assignment to investigate learning retardation among the school children of France over 60 years ago, educators and psychologists have been interested in, and concerned with problems of mental testing. During the first half of that period particularly, many intelligence tests were devised, a goodly proportion of which were designed to test the mentality of infants and very young children. At the present time a dozen or so of these tests are being widely used, and many others are in less frequent use by child psychologists and other clinicians. The purpose of the present project was to investigate and evaluate the total infant and preschool mental testing situation, and to arrive at some conclusions as to the current needs in this area.

Theories and Points of View. We made three approaches to our problem. First, it seemed important to trace the theoretical thinking that underlay the efforts of the makers of tests, and of those who were otherwise interested in the nature of intelligence and how it develops in children. Hence, a review of the literature dealing with these problems was made. We also reviewed the literature reporting research, and other uses that have been made of mental tests for infants and young children.

One aspect of the problem of intelligence concerning which different theoretical positions have been held is the question of the nature or composition (structure) of intelligence. The first attempts of note to appraise mental ability involved the testing of separate, discrete functions, mostly sensory and sensorimotor (Galton, 1869). Somewhat later, attempts were made to test the higher mental processes (Binet, 1905-1914); out of these approaches came two points of view: (1) Intelligence consists of many specific abilities not bound together by a common factor, and (2) intelligence is general mental ability. With the support of statistical analysis (Spearman, 1904-1927), the latter position became quite widely accepted. With the development of the factor-analytic methods, however, general intelligence (g) was found to break down into many different, but related abilities (factors).

A second theoretical problem has been that of the determinants of intelligence. Galton developed his program of eugenics on the theory that mental abilities are inherited just as are physical features. The general idea that intelligence is "fixed" genetically has had general acceptance in the past. The related doctrine of predeterminism held that, not only is one's level of intelligence fixed by the genes, but so also are the course, and the rate and timing of its development. The concept of a constant IQ is an expression of this predeterministic theory. This general position, which gives no weight to environment, was challenged by research findings allegedly showing marked effects of environment on IQ level. After a period of controversy among investigators, the current commonly accepted view is that intelligence develops through the interaction of organism and environment.

The problem of the nature of developmental change in mentality is a third area of intelligence theory. Does intelligence grow from infancy to maturity through simple accretion with no change in kind or quality, or is mentality in later childhood and adulthood different in quality and complexity than the mentality of the infant? Probably largely through the influence of Piaget (1952) the present trend of thinking is that intelligence develops in a qualitative, as well as a quantitative sense.

The positions assumed with regard to these three theoretical problem areas by students of intelligence have had much to do with the nature, the adequacy and general validity of the presently available mental tests, particularly of those at the infant, and early childhood levels. Most of the literature reporting the use of tests has been concerned with these problems in one way or another, or with further evaluation of the tests themselves.

Uses Currently Being Made of Infant and Preschool Mental Tests. Our second approach was to make a questionnaire survey of the uses being made of mental tests for young children. Our purposes were to determine the relative frequency with which each of the various tests are used, and for what specific purposes. We also asked for the judgments of the users as to the value of the tests, their affectiveness and limitations, and in what ways they felt the tests could be improved.

The questionnaire went out to 750 treatment agencies, research centers, and individual clinicians and researchers. Of the 330 returns, 113 indicated that children under six were not included in their clientele. Our results are based upon 217 completed returns.

Nearly 60 percent of the respondents represent educational institutions - nursery schools, kindergartens, day care centers, etc.

The appraisal of mental capacity was made for a variety of reasons, the most frequent (81%) of which was diagnosis of such conditions as mental retardation, speech or hearing defects, organic brain damage or other disabilities. Testing is also done in connection with treatment procedures. Types of treatment reported are psychotherapy (52%), parent counseling (9%), physical and occupational therapy (5%), and developmental guidance (3%).

Fifty-one different tests and scales were checked or mentioned as being in use with children under six years of age. The Stanford-Binet is by far most frequently used (90%). The six next most widely used tests in order of frequency are: Goodenough's Draw-a-man, the WISC, the Gesell Schedules, the Cattell Infant Scale, Ammons Picture-Vocabulary, and the Merrill-Palmer Scale.

Research activity reported was largely concerned with pathology or other abnormal conditions. Studies of autism and anoxia were reported by 13 percent and 8 percent respectively. Five percent reported developmental studies. About 2 percent reported studies to test the validity of tests in use. A number of other problems were reported, in each case, by only 1 or 2 people.

Special values, as well as inadequacies, were reported for each of the widely used tests. The most frequently mentioned limitations were: poor validity, manuals inadequate or difficult to use, limited norms, poor predictability, insufficient diagnostic precision, culturally outdated, too subjective, lacking in theoretical rationale as to dimensions measured.

Many suggestions were offered for the improvement of the available tests.

Factor Analyses of Infant and Preschool Mental Test Content. The third approach was through a statistical analysis of an accumulated mass of test results. Test protocols of 1,926 infants and young children were studied by factor analysis to determine the meaning content of the tests used. Item intercorrelations were obtained for 2 levels each of the Cattell Infant Scale, the California First Year Scale, and the Gesell Developmental Schedules, for 5 levels of the Merrill-Palmer Scale, and for 3 levels of the Stanford-Binet (Form L). Our data thus represented the unbroken course of developmental progress in mental ability from 3 months to 5 years. From each of these inter-item correlation matrices, principal axes were extracted. The factors were then rotated with three methods: varimax orthogonal, biquartimin oblique, and graphic orthogonal.

Considerable exploration was carried out to reduce the original test-item lists so as to eliminate from the analyses dependent variables and to give as nearly as possible, in each case, simple structure.

For each factor in the various analyses, the items having factor loadings of .40 or over were studied to discover the similarity of meaning that functioned to bring them together in the same factor. Suggestions as to factor names were made by three or more qualified judges and agreement reached as to probable meaning, in terms of Guilford's Structure of Intellect. The names for the factors are tabulated for all five test scales according to age level.

The general findings of these 14 factor analysis problems were as follows:

1. Even at the earliest ages, the nature of the factors indicated that the test responses may be definitely intellectual in nature. Whether or not the suggested names for the factors represent their true meaning, it is evident that children with wide individual variation, of course, even as young as three months of age, are capable of "thinking" processes.

2. The factor content of tests at different age levels for the same scale vary in content. Hence one of the reasons for longitudinal changes in mental test scores is that the child is being tested for different abilities at the two age levels.

3. The content in terms of tests, of the various scales also vary greatly. One scale will emphasize cognition and memory, with no items related to evaluation and production, while another scale emphasizes these abilities and neglects other processes.

4. Only the infant tests have items of divergent production, which are most closely associated with creativity. No wonder

intelligence tests have not been able to select the creative children when the tests do not have items testing this ability.

5. Some intelligence tests have items that are highly correlated, yielding some large, rather general factors. These general factors do not necessarily test "g", or general intelligence content, but rather are merely so narrow in their coverage of abilities, that there is little variations in their meaning.

6. The meaning of the test item is not necessarily related to the format and type of material used. Items containing blocks for manipulation may be just as effective an indication of cognition as are tests using words.

7. There is no evidence from this study that the results of the infant tests show less valid discrimination of ability than do the tests for older age levels. Some tests have a much wider range of test item meaning than others, but this is not necessarily a matter of age, but is due to the kind of items selected for the test scale.

8. Observational guessing about related item meaning is shown to have little validity in grouping together items in terms of the fundamental item-relationships, as shown by the factor analyses. For example, the Gesell items are quite differently classified by factor selectivity than they are by the labels in the Developmental Schedules.

General Conclusions. This project through its three-pronged approach, we feel, has demonstrated that there is a real need for more adequate means of appraising the mentality of young children. Recent research has provided a better understanding of the nature of intelligence, its determinants, and the nature of its developmental change, thus furnishing a sounder basis for new scale construction.

The users of presently available tests generally feel the need for improved testing devices and techniques. Our analyses of scale content have demonstrated a great lack of consistency among, and within the scales now in wide use in terms of factor content and meaning, thus pointing up the need for more consistent and adequate test scales.

Any attempt to develop a new intelligence scale, if it is to be an improvement over what is now available, must first and foremost be guided by, and be consistent with the best conceived, and most solidly based theory of the nature of mentality. Research has made it quite clear that the human being, even at very early ages, possesses not just one single general ability factor, but a number of abilities. It should not be assumed without proof however, that the same repertory of mental abilities will be found at all levels of development. Guilford's "structure of intellect" is presumed to be the structure to be found in young adults - the completed structure. The structure of the infantile intellect is in its early stages of development and therefore is presumably relatively simple, even though its constituent abilities are indeed "intellectual" in nature. Since the young child's intellect is developing, it is changing constantly by the emergence of new abilities.

However, as the repertory of abilities common to children at these different early age levels is determined through continued research, test items for their evaluation can be devised and appropriate scales constructed.

It must also be realized that it is always the level of acquired abilities that is available for testing, not the child's capacity or his mental potentiality. We can only observe or test what the child can do - indeed what he has "learned" to do. Children, of course, through use of their present abilities, acquire new abilities that are within the level of their developed capacity. A child who has been deprived of an environment which would stimulate and allow the exercise of his cognitive and perceptual abilities may perform at a relatively low level on a test, not because he has a low capacity, but simply because he has not learned to do the things required in the test-performance which may be well within his capacity to acquire. Perhaps one of the reasons for the lack of "predictive validity" of infant tests is that babies, due to wide variations in environmental opportunity and stimulation, have, to varying degrees, acquired the abilities that are within the limits of their developed capacity at the time of testing. Later on, perhaps as school children, their relative positions with respect to manifest ability will have changed.

References

Allport, F. H. *Social Psychology*. Cambridge: Houghton Mifflin, 1924.

Ammons, R. B., & Ammons, H. S. *The Full Range Vocabulary Test*. New Orleans: Authors, 1948.

Ammons, R. B., & Holmes, J. C. The Full-Range Picture Vocabulary Tests: III, Results for a Preschool Age Population. *Child Develop.*, 1949, 20, 5-14.

Anastasi, A., & De Jesus, C. Language Development and Non-Verbal IQ of Puerto Rican Preschool Children in New York City. *J. Abnor. & Social Psychol.*, 1953, 48, 357-366.

Anastasi, A. *Psychological Testing*. New York: Macmillan, 1961.

Anderson, J. E. The Prediction of Terminal Intelligence from Infant and Preschool Tests. 39th. *Yearb. Nat. Soc. Studies Educ.*, 1940, Part I, 385-403.

Anderson, L. D. The Predictive Value of Infancy Tests in Relation to Intelligence at Five Years. *Child Develop.*, 1939, 10, 203-212.

Arthur, G. A Point Scale of Performance Tests. Vol. I, *Clinical Manual*. New York: Commonwealth Fund, 1930.

Arthur, G. *A Point Scale of Performance Tests*. (Revised Form II). New York: Psychological Corp., 1947.

Arthur, G. *The Arthur Adaptation of the Leiter International Performance Scale*. Psychol. Serv. Center Press, 1952.

Ausubel, D. P. *Theory and Problems of Child Development*. New York: Grune & Straton, 1958.

Baker, C. T., Sontag, L. W., & Nelson, V. L. Individual and Group Differences in the Longitudinal Measurement of Change in Mental Ability. Fels Monograph Series, *Monogr. Soc. for Research in Child Develop.*, 1958, vol. XXIII, 68, No. 2.

Baller, W. R., & Charles, D. C. *The Psychology of Growth and Development*. New York: Holt, Rinehart & Winston, 1961.

Bayley, N. *The California First Year Mental Scale*. Berkeley: Univer. of Calif. Press, 1933, Pp. 24.

Bayley, N. Mental Growth During the First Three Years. *Genet. Psychol. Monogr.*, 1933, 14, 1-93.

Bayley, N., & Jones, H. E. Environmental Correlates of Mental and Motor Development: a Cumulative Study from Infancy to Six Years. *Child Develop.*, 1937, 8, 329-341.

Bayley, N. Some Increasing Parent-Child Similarities During the Growth of Children. *J. Educ. Psychol.*, 1954, 45, 1-21.

Bayley, N. On the Growth of Intelligence. *Amer. Psychologist*, 1955, 10, 805-818.

Benton, A. L. Influence of Incentives upon Intelligence Test Scores of School Children. *J. Genet. Psychol.*, 1936, 49, 494-496.

Binet, A., & Simon, T. Peut-on Enseigner la Parole aux Sourds-muets? *L'annee Psychologique*, 1909, Vol. XV, 373-396.

Binet, A., & Simon, T. *The Development of Intelligence in Children*. Trans. by Elizabeth S. Kite. Baltimore: Williams & Wilkins, 1916.

Black, I. S. The Use of the Stanford-Binet (1937 Revision) in a Group of Nursery School Children. *Child Develop.*, 1939, 10, 157-171.

Bowlby, J. *Maternal Care and Mental Health*. Geneva: World Health Organization, 1951.

Bradway, K. P. An Experimental Study of the Factors Associated with Stanford-Binet IQ Changes from the Preschool to the Junior High School. *J. Genet. Psychol.*, 1945, 66, 107-128.

Bradway, K. P. Predictive Value of Stanford-Binet Preschool Items. *J. Educ. Psychol.*, 1945, 36, 1-16.

Bradway, K. P., Thompson, C. W., & Craven, R. B. Preschool IQ's After Twenty-five Years. *J. Educ. Psychol.*, 1958, 49, 278-281.

Brenner, A., & Morse, N. Measurement of Children's Readiness for School. *Pap. Mich. Acad. Sci.*, 1956, 41, 333-340.

Brenner, A. A New Gestalt Test for Measuring Readiness for School. *Merrill-Palmer Quart.*, 1959, 6, 1-25.

Burt, C. Mental and Scholastic Tests. London: King, 1921.

Burt, C., Jones, E., Miller, E., & Moodie, W. How the Mind Works. New York: Appleton-Century-Crofts, 1934.

Burt, C. The Analysis of Temperament. Brit. J. Psychol., 1938, 17, 158-188.

Carr, H. A. Psychology: A Study of Mental Activity. New York: Longmans Green, 1926.

Cattell, P. The Measurement of Intelligence of Infants and Young Children. New York: The Psychological Corp., 1940.

Cattell, R. B. The IPAT Test of G: Culture Free, Scale I. Champaign, Ill.: Inst. Pers. Abil. Test, 1950.

Cattell, R. B. Factor Analysis: An Introduction and Manual for the Psychologist and Social Scientist. New York: Harper, 1952.

Cattell, R. B. Research Designs in Psychological Genetics with Special Reference to Multiple Variance. Amer. J. Hum. Genet., 1953, 5, 76-91.

Cattell, R. B., & Muerle, J. L. The 'Maxplane' Program for Factor Rotation to Oblique Simple Structure. Educ. & Psychol. Measure., 1960, 20, 569-589.

Cavanaugh, M. C., Cohen, I., Dunphy, D., Ringwell, E. A. & Goldberg, I. D. Prediction from Cattell Infant Intelligence Scale. J. Consult. Psychol., 1957, 21, 33-37.

Crissey, O. L. Mental Development as Related to Institutional Residence and Educational Achievement. Univer. of Iowa Stud. on Child Welfare, 1937, vol. 13, No. 1.

Cunningham, B. Infant IQ Rating Evaluated after an Interval of Seven Years. J. Exper. Educ., 1934, 3, 84-87.

Dashiell, J. F. Fundamentals of Objective Psychology. Cambridge: Houghton Mifflin, 1928.

Dearborn, W. F., & Rothney, J. W. Harvard Growth Study: Predicting the Child's Development. Cambridge, Mass.: Sci.-Art Pub., 1941.

DeForest, R. A Study of the Prognostic Value of the Merrill-Palmer Scale of Mental Tests and the Minnesota Preschool Scale. J. Genet. Psychol., 1941, 59, 219-223.

Dennis, W., & Najarian, P. Infant Development under Environmental Handicap in Morris L. Haimowitz and N. R. Haimowitz (Ed.) Human Development. New York: Thomas Y. Crowell, 1960, Pp. 173-189.

Despert, J. L., & Pierce, H. O. The Relation of Emotional Adjustment to Intellectual Function. Genet. Psychol. Monogr., 1946, 34, 3-56.

Driscoll, G. P. The Development Status of the Preschool Child as a Prognosis of Future Development. Child Develop. Monogr. Teachers College: Columbia Univer., 1933, No. 13.

Ebert, E., & Simmons, K. The Brush Foundation Study of Child Growth and Development. I. Psychometric Tests. Monogr. Soc. Res. in Child Develop., 1943, 8, No. 2 (Whole No. 35).

Escalona, S. K. The Predictive Value of Psychological Tests in Infancy: A Report on Clinical Findings. Amer. Psychologist, 1948, 3, 281.

Escalona, S. K. The Use of Infant Tests for Predictive Purposes. Bull. Menninger Clin., 1950, 14, 117-128.

Escalona, S. K., & Moriarty, A. Prediction of Schoolage Intelligence from Infant Tests. Child Develop., 1961, 32, 597-605.

Frandsen, A., & Barlow, F. P. Influence of the Nursery School on Mental Growth. 39th. Yearb. Nat. Soc. Educ., 1940, Part II, 143-148.

French, J. W. The Factorial Invariance of Power-Factor Tests. J. Educ. Psychol., 1957, 48, 93-100.

Furfey, P. H. The Relation Between Socio-Economic Status and the Intelligence of Young Infants as Measured by the Linfert-Hierholzer Scale. J. Genet. Psychol., 1928, 35, 478-480.

Furfey, P. H., & Muehlenbein, J. The Validity of Infant Intelligence Tests. J. Genet. Psychol., 1932, 40, 219-223.

Gallaghar, J. J. Clinical Judgment and the Cattell Intelligence Scale. J. Consult. Psychol., 1953, 17, 303-305.

Galton, F. Hereditary Genius. New York: Appleton-Century-Crofts, 1869 and 1880 (reprint).

Galton, F. Inquiries into the Human Faculty and its Development. New York: Macmillan, 1883.

Garrett, H. E. A Developmental Theory of Intelligence. Amer. Psychologist, 1946, 1, 372-378.

Gesell, A. The Mental Growth of the Preschool Child. New York: Macmillan, 1925.

Gesell, A. Infancy and Human Growth. New York: Macmillan, 1928.

Gesell, A. The Ontogenesis of Infant Behavior. In D. Carmichael (Ed.), Manual of Child Psychology. New York: John Wiley, 1954.

Gilliland, A. R. The Measurement of the Mentality of Infants. Child Develop., 1948, 19, 155-158.

Goddard, H. H. The Binet and Simon Tests of Intelligence Capacity. Training Sch., 1908.

Goddard, H. H. The Binet-Simon Measuring Scale for Intelligence. (Revised Edition). Training Sch., 1911.

Goldfarb, W. Infant Rearing and Problem Behavior. Amer. J. Orthopsychiat., 1943, 13, 249-265.

Goldfarb, W. Variations in Adolescent Adjustment of Institutionally Reared Children. Amer. J. Orthopsychiat., 1947, 17, 449-457.

Goodenough, F. L. Racial Differences in the Intelligence of School Children. J. Exp. Psychol., 1926, 9, 388-397.

Goodenough, F. L. A Preliminary Report on the Effects of Nursery School Training upon Intelligence Test Scores of Young Children. 27th. Yearb. Nat. Soc. Educ., 1928a, 361-369.

Goodenough, F. L. The Relation of Intelligence of Preschool Children to the Occupation of their Fathers. Amer. J. Psychol., 1928, 40, 284-302.

Goodenough, F. L., Maurer, K. M., & Van Wagenen, M. J. Minnesota Preschool Scales. Minneapolis: Educ. Test Bur., 1932-1940.

Goodenough, F. L. Look to the Evidence: A Critique of Recent Experiments on Raising the IQ. Edu. Methods, 1939, 19, 73-79.

Goodenough, F. L., & Maurer, K. The Mental Development of Nursery School Children Compared with That of Non-Nursery School Children. 29th. Yearb. Nat. Soc. Educ., 1940, Part II, 161-178.

Goodenough, F. L. New Evidence on Environmental Influence on Intelligence. 39th. Yearb. Nat. Soc. Educ., 1940, Part I, 307-365.

Goodenough, F. L. Studies of the 1937 Revision of the Stanford-Binet Scale: I. Variability of the IQ at Successive Age-Levels. J. Educ. Psychol., 1942, 33, 241-251.

Grassi, J. R. School Psychological Clinics, Part III: the Preschool Psychological Clinic. North Carolina Med. J., 1955, 16, 171-180.

Griffiths, R. The Abilities of Babies. London: Univer. of London Press, 1954.

Guilford, J. P. Factor Analysis in a Test Development Program. Psychol. Rev., 1948, 55, 79-94.

Guilford, J. P. The Structure of Intellect. Psychol. Bull., 1956, 53, 267-293.

Guilford, J. P. A Revised Structure of Intellect. Rep. Psychol. Lab. Los Angeles: Univer. Southern Calif., 1957, No. 19.

Guilford, J. P. A System of Psychomotor Abilities. Amer. J. Psychol., 1958, 71, 164-174.

Guilford, J. P. Three Faces of Intellect. Amer. Psychologist, 1959, 14, 469-479.

Guilford, J. P., Merrifield, P. R., Christensen, P. R., & Frick, J. W. A Factor-Analytic Study of Problem-Solving Abilities. Rep. Psychol. Lab. Los Angeles: Univer. Southern Calif., 1960, No. 22.

Guilford, J. P., & Merrifield, P. R. The Structure of Intellect Model: Its Uses and Implications. Rep. Psychol. Lab. Los Angeles: Univer. Southern Calif., 1960, No. 24.

Harman, H. H. Modern Factor Analysis. Chicago: Univer. of Chicago, 1960.

Harris, D. H. Measuring the Psychological Maturity of the Child: a Revision and Extension of the Goodenough Draw-a-Man Test. New York: World Book, 1961.

Hastings, H. J. The Predictive Value of Individual Items in Preschool Intelligence Tests. Unpublished doctoral thesis, Univer. of Calif., 1952.

Hilden, A. H. A Longitudinal Study of Intellectual Development. J. Psychol., 1949a, 28, 187-214.

Hofstaetter, P. R. The Changing Composition of "Intelligence," A Study in T Technique. *J. Genet. Psychol.*, 1954, 85, 159-164.

Honzik, M. P., Macfarlane, J. W., & Allen, L. The Stability of Mental Test Performance between Two and Eighteen Years. *J. Exp. Educ.*, 1948, 17, 309-324.

Hotelling, H. Analysis of a Complex of Statistical Variables into Principal Components. *J. Educ. Psychol.*, 1933, 24, 417-520.

Humphrey, L. G. The Organization of Human Abilities. *Amer. Psychologist*, 1962, 17, 475-483.

Hunt, J. M. *Intelligence and Experience*. New York: Ronald Press, 1961.

Hurst, J. G. A New Factor Analysis of the Merrill-Palmer with Reference to Theory and Test Construction. *Educ. Psychol. Measmt.*, 1960, 20, 519-532.

Hurst, J. G. The Theoretical and Methodological Problems Arising from the Application of Factor Analytic Techniques to Existing Preschool Intelligence Scales. Paper read at Amer. Psychol. Ass., New York, 1961.

Illingworth, R. S. *The Development of the Infant and Young Child. London: Livingston, 1960.*

Illingworth, R. S. The Predictive Value of Developmental Tests in the First Year with Special Reference to the Diagnosis of Mental Subnormality. *J. Child Psychol. & Psychiat.*, 1961, 2, 210-215.

Jones, H. E., & Seashore, R. H. The Development of Fine Motor and Mechanical Abilities. In N. B. Henry (Ed.), *Yearb. Nat. Soc. Stud. Educ.*, 1944, 43, Part I.

Jones, L. V. A Factor Analysis of the Stanford-Binet at Four Age Levels. *Psychometrika*, 1949, No. 4, 299-331.

Jones, L. V. Primary Abilities in the Stanford-Binet Age 13. *J. Genet. Psychol.*, 1954, 84, 125-147.

Kawin, E., & Hoefer, C. *A Comparative Study of a Nursery School Versus a Non-Nursery School Group*. Chicago: Univer. of Chicago Press, 1931.

Kawin, E. Children of Preschool Age. Chicago: Univer. of Chicago Press, 1934.

Kelley, T. L. Essential Traits of Mental Life. Cambridge: Harvard Univer. Press, 1935.

Klatskin, E. H. Intelligence Test Performance at One Year among Infants Raised with Flexible Methodology. J. Clin. Psychol., 1952, 8, 230-237.

Knobloch, H., & Pasamanick, B. The Distribution of Intellectual Potential in an Infant Population. In The Epidemiology of Mental Disorder. Amer. Ass. for the Advancement of Science Publication, 1959, No. 60, 249-272.

Knobloch, H., & Pasamanick, B. Environmental Factors Affecting Human Development: Before and After Birth. Pediatrics, 1960, 26, 210-218.

Knobloch, H., & Pasamanick, B. An Evaluation of the Consistency and Predictive Value of the Forty Week Gesell Development Schedule. In C. Shagass and B. Pasamanick (Eds.), Child Develop. & Child Psychiatry. Psychiatric Research Report of the Amer. Psychiat. Ass., 1960, No. 13, 10-31.

Koch, M. B., & Meyer, D. R. A Relationship of Mental Age to Learning-Set Formation in the Preschool Child. J. Comp. Physiol. Psychol., 1959, 52, 387-389.

Kohs, S. C. Intelligence Measurement: A Psychological and Statistical Study Based upon the Block-Design Test. New York: Macmillan, 1923.

Kuhlmann, F. A Revision of Binet-Simon System for Measuring the Intelligence of Children. J. Psychol. Asthenics. (Monogr. Supplmt.), 1912.

Kuhlmann, F. A Handbook of Mental Tests. Baltimore: Warwick & York, 1922.

Leiter, R. G. International Performance Scale. Washington, D. C.: Psychol. Serv. Center Press, 1948.

Leiter, R. G. The Leiter International Performance Test. Calif.: Author, 1950.

Leiter, R. G. Part I of the Manual for the 1948 Revision of the Leiter International Performance Scale. Psychol. Serv. Cent. J., 1959, 11, 1-72.

Levinson, B. M. A Comparison of the Performance of Bilingual and Monolingual Native Born Jewish Preschool Children in Traditional Parentage on Four Intelligence Tests. *J. Clin. Psychol.*, 1959, 15, 74-76.

Linfert, H. E., & Hierholzer, H. M. A Scale for Measuring the Mental Development of Infants During the First Year of Life. *Stud. Psychol. Psychiat.*, Catholic Univer. Amer., 1928, 1, No. 4.

Maurer, K. M. *Intellectual Status at Maturity as a Criterion for Selecting Items in Preschool Tests.* Minneapolis: Univer. of Minn. Press, 1946.

McCarthy, D. A Study of Reliability of Goodenough Drawing Test of Intelligence. *J. Psychol.*, 1944, 18, 201-216.

McCarthy, D. Measurement of Cognitive Abilities at the Preschool and Early Childhood Levels. In *Proc. 1958 Invitat. Conf. Test Prob.* Educational Testing Serv., 1959, Pp. 10-25.

McHugh, G. Changes in IQ at the Public School Kindergarten Level. *Psychol. Monogr.*, 1943, 55, No. 2.

Merrifield, P. R., & Cliff, N. Factor Analytic Methodology. *Rev. Educ. Research* (Statistical Methodology), 1963, vol. 33, No. 5.

Merrifield, P. R., Guilford, J. P., & Gershon, A. *Differentiation of Divergent-Production Abilities at the Sixth Grade Level.* Report from Psychological Lab., Calif.: Dept. of Psychology, Univer. Southern Calif., 1963, No. 27.

Meyers, C. E., & Dingman, H. F. The Structure of Abilities at the Preschool Ages: Hypothesized Domain. *Psychol. Bull.*, 1960, 57, 514-532.

Meyers, C. E., Orpit, R. E., Attwell, A. A., & Dingman, H. F. Primary Abilities at Mental Age Six. *Mon. Soc. Res. in Child Develop.*, 1962, 27, 1-40.

Moriarty, A. Children's Ways of Coping with the Intelligence Test. *Bull. Menninger Clin.*, 1960, 24, 115-127.

Mowrer, W. M. C. Intelligence Scale for Preschool Children. *Child Develop.*, 1933, 4, 318-322.

Mowrer, W. M. C. Performance of Children in Stutsman Tests. *Child Develop.*, 1934, 5, 93-96.

Munn, N. L. *Psychology*. Cambridge: Houghton Mifflin, 1946.

Nelson, V. L., & Richards, T. W. Studies in Mental Development: I. Performance on Gesell Items at Six Months and its Predictive Value for Performance on Mental Tests at Two and Three Years. *J. Genet. Psychol*., 1938, 52, 303-325.

Nelson, V. L., & Richards, T. W. Studies in Mental Development: III. Performance of Twelve Months Old Children on the Gesell Schedule and its Predictive Value for Mental Status at Two and Three Years. *J. Genet. Psychol*., 1939, 54, 181-191.

Pease, D., Rosauer, J. K., & Wolins, I. Reliability of Three Infant Developmental Scales Administered during the First Year of Life. *J. Genet. Psychol*., 1961, 98, 295-298.

Peters, C. C., & McElwee, A. R. Improving Functional Intelligence by Analytical Training in Nursery School. *Elem. Sch. J.*, 1944, 45, 213-219.

Peterson, I. J. A Preliminary Study of the Effects of Previous Nursery School Attendance upon the First-Year Children Entering Kindergartens. *Stud. in Preschool Educ*., Iowa City: Univer. of Iowa, 1937, 197-248.

Piaget, J. 1936. *The Origins of Intelligence in Children*. Trans. by Margaret Cook. New York: Int. Univer. Press, 1952.

Pinnean, S. R. A Critique of the Article by Margaret Ribble. *Child Develop*., 1950, 21, 203-228.

Pinnean, S. R. *Changes in the Intelligence Quotient: Infancy to Maturity*. Boston: Houghton Mifflin, 1961.

Pintner, R., & Paterson, D. G. *A Scale of Performance Tests*. New York: Appleton, 1917.

Porteus, S. D. *Guide to Maze Test*. Vineland, New Jersey: The Training School, 1924.

Richards, T. W., & Nelson, V. L. Studies in Mental Development: II. Analysis of Abilities Tested at the Age of Six Months by the Gesell Schedules. *J. Genet. Psychol*., 1938, 52, 327-331.

Richards, T. W., & Nelson, V. L. Abilities of Infants During the First Eighteen Months. *J. Genet. Psychol*., 1939, 55, 299-318.

Starkweather, E., & Roberts, K. IQ Changes Occuring during Nursery Attendance at the Merrill-Palmer School. 39th. *Yearb. Nat. Soc. Stud. Educ*., 1940, Part II, 315-336.

Schafer, S., & Leitch, M. An Exploratory Study of the Usefulness of a Battery of Psychological Tests with Nursery School Children. Amer. J. of Psychiat., 1948, 104, 647-652.

Shotwell, A., & Gilliland, A. R. A Preliminary Scale for the Measurement of the Mentality of Infants. Child Develop., 1943, 14, 167-177.

Simon, A. J., & Bass, L. G. Toward a Validation of Infant Testing. Amer. J. Orthopsychiat., 1956, 26, 340-350.

Simpson, B. R. The Wandering IQ: Is It Time to Settle Down? J. Psychol., 1939, 7, 351-367.

Skeels, H. M., & Filmore, E. A. Mental Development of Children from Underprivileged Homes. Ped. Sem. & J. Genet. Psychol., 1937, 50, 427-439.

Skeels, H. M. Mental Development of Children in Foster Homes. J. Consult. Psychol., 1938, No. 2, 33-43.

Skeels, H. M., Updegraff, R., Wellman, B. L., & Williams, H. M. A Study of Environmental Stimulation: an Orphanage Preschool Project. Univer. Iowa Stud. Child Welf., 1938, 15, No. 4.

Skodah, M. Children in Foster Homes: A Study of Mental Development. Univer. Iowa Stud. Child Welf., 1939, 16, No. 1.

Spearman, C. General Intelligence Objectively Determined and Measured. Amer. J. Psychol., 1904, 15, 201-293.

Spearman, C. The Theory of Two Factors. Psychol. Rev., 1914, 21, 101-115.

Spearman, C. The Abilities of Man. New York: Macmillan, 1927.

Spiker, C. C., & Irwin, O. C. The Relationship Between IQ and Indices of Infant Speech Sound Development. J. Speech & Hearing Disorders, 1949, 14, 335-343.

Spitz, R. A. The Role of Ecological Factors in Emotional Development in Infancy. Child Develop., 1949, 20, 145-155.

Stern, W. L. The Psychological Methods of Testing Intelligence. Trans. by G. M. Whipple. Baltimore: Warwick & York, 1914.

Stoddard, G. D., & Wellman, B. L. Child Psychology, New York: Macmillan, 1934.

Stoddard, G. D., & Wellman, B. L. Environment and the IQ. 39th. Yearb. Nat. Soc. Stud. Educ., 1940, Part I, 405-442.

Stoddard, G. D. The Meaning of Intelligence. New York: Macmillan, 1943.

Stone, L. J. A Critique of Studies of Infant Isolation. Child Develop., 1954, 25, 9-20.

Stutsman, R. Mental Measurement of Preschool Children. New York: World Book, 1931.

Tate, M. E. The Influence of Cultural Factors on the Leiter International Performance Scale. J. Abnorm. Soc. Psychol., 1952, 47, 497-501.

Teagarden, F. M. Merrill-Palmer Scale of Mental Tests. Mental Measmt. Yearb., 1940, 230.

Terman, L. M. The Measurement of Intelligence. Boston: Houghton Mifflin, 1916.

Terman, L. M., & Merrill, M. A. Measuring Intelligence. Boston: Houghton Mifflin, 1937.

Thorndike, E. L. Educational Psychology. New York: Columbia Univer., 1914, Vol. III.

Thurstone, L. L. The Vectors of the Mind. Chicago: Univer. of Chicago Press, 1935.

Thurstone, L. L. Primary Mental Abilities. Chicago: Univer. of Chicago Press, 1938.

Thurstone, L. L., & Thurstone, T. G. Factorial Studies of Intelligence. Psychometr. Monogr., 1941, No. 2.

Thurstone, L. L. Psychological Implications of Factor Analysis. Amer. Psychologist, 1948, 3, 402-408.

Thurstone, L. L., & Thurstone, T. G. SRA Primary Abilities for ages 5-7. Sci. Res. Associates, 1953.

Trabue, M. R., & Stockbridge, F. P. Measure Your Mind. New York: Doubleday Page, 1920.

Van Alstyne, D. The Environment of Three-Year-Old Children: Factors Related to Intelligence and Vocabulary Tests. Teach. Coll. Columbia Univer., Contrib. Educ., 1929, No. 366.

Wellman, B. L. Some New Bases for Interpretation of the IQ. J. Genet. Psychol., 1932a, 41, 116-126.

Wellman, B. L. The Effects of Preschool Attendance upon the IQ. J. Exp. Educ., 1932b, 1, 48-69.

Wellman, B. L. Growth in Intelligence under Different School Environments. J. Exp. Educ., 1934, 3, 59-83.

Wellman, B. L., & Coffey, H. S. The Role of Cultural Status in Intelligence Changes for Preschool Children. J. Exp. Educ., 1936, 5, 191-202.

Wellman, B. L. Mental Growth from the Preschool to College. J. Exp. Educ., 1937, 6, 127-138.

Wellman, B. L. The Intelligence of Preschool Children as Measured by the Merrill-Palmer Scale of Performance Tests. Univer. Iowa Stud. on Child Welf., 1938, 15, No. 3.

Wright, R. E. A Factor Analysis of the Original Stanford-Binet Scale. Psychometrika, 1939, 4, 209-220.

Yerkes, R. M., Bridges, J. W., & Hardwick, R. S. A Point Scale for Measuring Mental Ability. Baltimore: Warwick & York, 1915.

Yerkes, R. M., & Foster, J. C. 1923 Revision: A Point Scale for Measuring Mental Ability. Baltimore: Warwick & York, 1923.

APPENDIX

Table 1. TYPES OF SERVICES OFFERED CHILDREN AND THEIR FAMILIES, WITH PERCENTAGE FREQUENCIES

Type of Service	Percent Reporting
Testing	
Intelligence	93
As a diagnostic procedure	81
Pre-adoption	36
Pre-admission	32
Children with physical handicaps	40
Children with speech and hearing disabilities	3
Behavior and personality problems	2
Organic brain damaged	2
Other disabilities	1
Treatment	
Psychotherapy	52
Remedial (speech and hearing, reading, etc.)	21
Parent counseling	9
Physical and occupational therapy	5
Developmental guidance	3
Other	3
Educational	
Nursery schools	28
Residential schools	9
Kindergartens	8
Day care centers	8
Work with educationally handicapped	5
Other	2
Research Problems Investigated	
Autism	13
Anoxia	8
General developmental and longitudinal studies	5
Validation or infant tests	2
Adoptions	2
Speech and hearing problems	2
Study of hyperkinesis	2
Hereditary factors	2
Perceptual disorders	1
Effects of training	1
Epilepsy	1
Prematurity	1
Neurological impairment	1
Other problems, such as mother-infant interaction, cognitive development, creativity, cerebral palsy, aphasia, etc.	8

Table 1 (Continued)

Type of Service	Percent Reporting
Medical Services	
Psychiatric	31
Outpatient	25
Pediatric	22
Neurological	5
General physical	1
Diagnostic	1
Orthopedic	1
Electroencephalography	1

Table 2. TYPES OF CASES, MOST FREQUENTLY REPORTED BY 217 RESPONDENTS

Case Category	Percentage Frequency
Mentally retarded	76
Behavior Difficulties	75
Normal children	74
Brain Damaged	69
Physically Handicapped or Disabled	64
Verbal Difficulties	50
Autism	45
Psychotic children	38
Deaf or Blind children	3
Neurotic children	3
"Personality" Problems	2
Gifted children	2

Table 3. THE FIFTEEN MOST FREQUENTLY USED INFANT AND PRESCHOOL AGE MENTAL TEST SCALES, AS REPORTED BY 217 RESPONDENTS

Test Name	Percentage Frequency of Use
Stanford-Binet	90
Goodenough Draw-a-Man	73
WISC	54
Cattell Infant Scale	44
Gesell Developmental Schedules	36
Ammons Picture Vocabulary Test	33
The Merrill-Palmer Scale	28
Columbia	21
Grace-Arthur Performance	20
Minnesota Pre-School Scale	15
Raven Matrices	15
Leiter International	12
Kuhlman-Binet	7
California Pre-school	4
Griffith's Abilities of Babies Scale	3

Table 4. MENTAL TESTS, WITH AGE LEVELS, USED IN THE FOURTEEN FACTOR ANALYSIS STUDIES

Test Scale	Age-Level Analyzed (Months)	No. of Cases	Table No.
Cattell Infant Scale	3	150	5
	6	176	6
California First-Year Mental Scale	6	134	7
	12	137	8
Gesell Developmental Schedules	6	104	9
	12	104	10
Merrill-Palmer Scale	24	215	11
	30	142	12
	36	188	13
	48	166	14
	54	150	15
Stanford-Binet Scale	36	110	16
	48	76	17
	60	84	18
Totals	14 studies	1,926 Test Protocols	

TABLE 5. FACTOR ANALYSIS, CATTELL INFANT SCALE
AGE LEVEL, 3 MONTHS

Item Name	Item Level	Test Item Intercorrelations								Principal Axes (Unrotated)				h^2
		3-1	3-2	3-3	3-4	3-5	4-1	4-2	4-5	1	2	3	4	
Follow ring, circ. motion	3-1		835	905	895	807	651	633	740	.905	.319	.039	.167	.950
Anticipates feeding	3-2	835		680	746	780	782	620	754	.853	.002	-.157	.314	.850
Regards cube	3-3	905	680		896	795	627	491	763	.863	.420	.024	-.126	.937
Regards spoon	3-4	895	746	896		905	748	734	829	.946	.153	.111	-.131	.948
Inspects fingers	3-5	807	780	795	905		929	691	783	.941	-.050	-.234	-.173	.972
Finger manipulation	4-1	651	782	627	748	929		805	808	.890	-.381	-.199	-.057	.980
Hands remain open	4-2	633	620	491	734	691	805		807	.790	-.387	.275	.026	.850
Inc. act. at sight of toy	4-5	740	754	763	829	783	808	807		.891	-.119	.171	.008	.837
									LAMBDA	6.279	.613	.239	.194	
									PERCENT	78.49	7.66	2.98	2.42	

Item Name	Item Level	Varimax Rotation				Graphic Rotation (Orthogonal)				Biquartimin Rotation			
		Factors				Factors				Factors			
		1	2	3	4	A	B	C	D	1	2	3	4
Follow ring, circ. motion	3-1	.776	-.320	-.202	.453	.94	.05	.04	.24	.779	-.075	-.081	-.295
Anticipates feeding	3-2	.451	-.361	-.376	.612	.74	.25	.26	.41	.256	-.130	.193	-.545
Regards cube	3-3	.884	-.221	-.271	.184	.96	-.09	.07	-.04	.999	.056	.035	.034
Regards spoon	3-4	.758	-.475	-.344	.172	.95	.20	.08	-.06	.751	-.294	.090	.076
Inspects fingers	3-5	.563	-.378	-.679	.226	.83	.26	.45	-.04	.422	-.101	.582	-.042
Finger manipulation	4-1	.271	-.569	-.705	.291	.65	.56	.49	.06	-.037	-.411	.629	-.145
Hands remain open	4-2	.258	-.814	-.290	.194	.61	.70	.03	.03	.008	-.887	.054	.007
Inc. act. at sight of toy	4-5	.505	-.645	-.320	.251	.79	.45	.08	.05	.357	-.579	.069	-.040
	LAMBDA	2.873	2.053	1.519	[illegible]79								
	PERCENT	35.91	25.66	18[illegible]	[illegible]99								

TABLE 6. FACTOR ANALYSIS, CATTELL INFANT SCALE
AGE LEVEL, 6 MONTHS

Item Name	Item Level	Test Item Intercorrelations								Principal Axes (Unrotated)				h^2
		5-3	6-1	6-3	6-4	6-5	7-1	7-3	7-5	1	2	3	4	
Trans. obj. hand to hand	5-3		852	689	726	712	831	605	722	.871	-.015	-.281	-.246	.899
Secure cube on sight	6-1	852		757	849	809	809	724	724	.927	.088	-.120	-.020	.881
Fingers refl. in mirror	6-3	689	757		785	654	791	689	700	.844	-.030	-.072	.172	.748
Reaches unilaterally	6-4	726	849	785		814	658	647	551	.865	.393	-.095	.248	.973
Reaches persistently	6-5	712	809	654	814		836	731	519	.868	.286	.273	-.147	.930
Attempts pick up pellet	7-1	831	809	791	658	836		793	715	.913	-.108	.114	-.213	.904
Inspects ring	7-3	605	724	689	647	731	793		777	.838	-.237	.290	.137	.860
Exploits paper	7-5	722	724	700	551	519	715	777		.795	-.421	-.106	.101	.830
									LAMBDA	5.997	.489	.290	.248	
									PERCENT	74.97	6.12	3.62	3.10	

Item Name	Item Level	Varimax Rotation				Graphic Rotation (Orthogonal)				Biquartimin Rotation			
		Factors				Factors				Factors			
		1	2	3	4	A	B	C	D	1	2	3	4
Trans. obj. hand to hand	5-3	.369	-.366	.299	-.734	.71	.45	-.05	.41	.163	.082	.171	-.700
Secure cube on sight	6-1	.547	-.426	.374	-.511	.84	.31	.08	.22	.421	.207	.159	-.345
Fingers refl. in mirror	6-3	.525	-.544	.257	-.332	.78	.34	.10	-.01	.448	.483	-.044	-.081
Reaches unilaterally	6-4	.845	-.270	.331	-.278	.98	-.04	.01	.05	.929	.015	.044	-.035
Reaches persistently	6-5	.478	-.252	.740	-.302	.78	.04	.42	.36	.279	-.013	.699	-.147
Attempts pick up pellet	7-1	.258	-.521	.569	-.492	.67	.45	.35	.32	-.060	.385	.466	-.349
Inspects ring	7-3	.290	-.726	.479	-.139	.67	.48	.43	-.01	.054	.844	.227	.185
Exploits paper	7-5	.227	-.773	.136	-.403	.58	.68	.13	-.06	.019	.881	-.181	-.163
	LAMBDA	1.855	2.143	1.523	1.504								
	PERCENT	23.19	26.78	19.04	18.79								

TABLE 7. FACTOR ANALYSIS, CALIFORNIA FIRST YEAR MENTAL SCALE
AGE LEVEL, 6 MONTHS

Item Name	Item No.	Test Item Intercorrelations											Principal Axes (Unrotated)						h²
		47	49	50	51	55	56	58	59	60	61	63	1	2	3	4	5	6	
Accepts second cube	47		079	655	196	328	367	360	362	469	429	103	-.555	.114	.054	.293	.194	.290	.531
Vocal displeasure	49	079		218	258	512	239	426	141	411	252	059	-.430	-.160	-.575	-.364	-.235	.097	.738
Reaches persist.	50	655	218		662	578	634	585	482	673	589	523	-.904	.302	.160	-.051	.204	.255	1.000
Turns after spoon	51	196	258	662		487	332	394	605	207	558	084	-.666	-.328	.396	-.493	.149	-.031	.975
Bangs in play	55	328	512	578	487		532	328	444	427	425	064	-.651	-.224	-.158	-.157	-.057	.247	.589
Sust. insp. of ring	56	367	239	634	332	532		617	567	698	617	123	-.774	-.045	-.139	.258	.018	-.189	.723
Vocal satisfaction	58	360	426	585	394	328	617		390	497	262	331	-.675	.201	-.331	-.137	.330	-.331	.843
Lift cup with handle	59	362	141	482	605	444	567	390		533	666	-005	-.708	-.347	.235	.157	-.019	-.211	.746
Exp. string play	60	469	411	673	207	427	698	497	533		571	158	-.758	.034	-.264	.333	-.104	.023	.768
Rotates wrist	61	429	252	589	558	425	617	262	666	571		358	-.772	-.040	.296	.105	-.422	-.057	.877
Smiles at mirror image	63	103	059	523	084	064	123	331	-005	158	358		-.324	.760	.146	-.246	-.224	-.089	.822
												LAMBDA	5.011	1.030	.904	.780	.509	.422	
												PERCENT	45.55	9.36	8.22	7.09	4.63	3.83	

Item Name	Item No.	Varimax Rotation						Graphic Rotation (Orthogonal)					Biquartimin Rotation					
		Factors						Factors					Factors					
		1	2	3	4	5	6	A	B	C	D	E	1	2	3	4	5	6
Accepts second cube	47	-.066	.037	-.019	.679	-.225	-.117	-.03	.23	.61	.06	.10	-.054	.012	-.060	.715	.069	-.018
Vocalizes displeasure	49	-.073	.035	-.832	.005	-.085	-.178	.01	-.13	.19	.80	.21	.001	.004	.862	-.096	-.018	-.079
Reaches persistently	50	-.400	.398	-.155	.756	-.249	-.261	.19	.57	.67	.12	.38	.262	-.327	.046	.733	.028	-.124
Turns after fallen spoon	51	-.921	.058	-.167	.137	-.248	-.125	.62	.06	.30	.06	.70	.877	.007	.099	.023	.166	-.062
Bangs in play	55	-.321	-.030	-.530	.360	-.261	-.077	.23	-.04	.46	.40	.30	.209	.094	.500	.299	.121	.051
Sustained inspection of ring	56	-.077	.018	-.199	.364	-.587	-.448	.08	.10	.78	.24	.10	-.097	.042	.090	.216	.528	-.355
Vocalizes satisfaction	58	-.172	.201	-.258	.234	-.142	-.795	-.21	.30	.51	.34	.46	.083	-.155	.165	.103	.007	-.763
Lifts cup with handle	59	-.414	-.126	-.034	.213	-.683	-.215	.42	-.08	.68	.00	.19	.268	.188	-.075	.050	.694	-.139
Exploitive string play	60	.102	.056	-.335	.488	-.555	-.309	-.01	.11	.80	.33	-.06	-.295	.003	.248	.383	.463	-.179
Rotates wrist	61	-.288	.296	-.160	.248	-.786	.049	.62	.21	.62	.20	-.10	.111	-.238	.068	.074	.813	.187
Smiles at mirror image	63	-.020	.890	-.010	.106	-.059	-.119	-.05	.75	.04	.46	-.06	-.016	-.890	-.031	.042	.026	-.069
LAMBDA		1.420	1.105	1.271	1.723	2.010	1.126											
PERCENT		12.91	10.05	11.55	15.67	18.27	10.23											

TABLE 8. FACTOR ANALYSIS, CALIFORNIA FIRST YEAR MENTAL SCALE
AGE LEVEL, 12 MONTHS

Item Name	Item No.	Test Item Intercorrelations													Principal Axes (Unrotated)						
		80	84	88	90	91	92	94	95	96	97	99	100	102	1	2	3	4	5	6	h^2
Attempt sec. 3 blocks	80		718	570	642	491	182	328	446	597	145	469	168	287	.76	.28	-.07	-.24	-.23	-.03	.77
Pk. up cup, secure cube	84	718		554	738	525	169	294	543	411	215	282	307	050	.73	.54	.05	-.02	.09	.01	.84
Cube in cup on command	88	570	554		532	296	029	407	599	416	180	167	365	236	.65	.25	-.16	.26	-.10	.04	.59
Unwraps cube	90	642	738	532		562	180	376	460	402	272	348	142	190	.72	.33	.09	-.16	.04	-.10	.66
Holds crayon adaptively	91	491	525	296	562		459	298	281	580	448	246	173	225	.66	.05	.45	-.21	.09	.18	.73
Inhib. on command	92	182	169	029	180	459		259	146	162	296	134	123	465	.35	-.33	.35	-.23	.13	.19	.46
Strikes doll in imitat.	94	328	294	407	376	298	259		154	512	460	400	294	303	.59	-.23	.28	.28	-.19	-.46	.81
Imitates words	95	446	543	599	460	281	146	154		450	375	596	358	367	.72	.02	-.50	.09	-.36	.13	.92
Stirs with spn. in imit.	96	597	411	416	402	580	162	512	450		190	342	469	479	.72	-.10	.04	.11	.37	.16	.70
Holds cup to drink	97	145	215	180	272	448	296	460	375	190		279	339	236	.47	-.26	.28	.22	.44	-.12	.62
Says two words	99	469	282	167	348	246	134	400	596	342	279		136	573	.59	-.32	-.37	-.28	.09	-.34	.79
Dangles ring by string	100	168	307	365	142	173	123	294	358	469	339	136		319	.45	-.16	-.03	.47	-.26	.23	.50
Uses expressive jargon	102	287	050	236	190	225	465	303	367	479	236	573	319		.51	-.69	-.18	-.17	-.17	.18	.87
													% of Variance		39.0	11.0	7.2	5.6	4.8	4.3	

Item Name	Item No.	Varimax Rotation						Graphic Rotation (Orthogonal)						Biquartimin Rotation					
		Factors						Factors						Factors					
		1	2	3	4	5	6	A	B	C	D	E	F	1	2	3	4	5	6
Attempt sec. 3 blocks	80	.781	.276	.154	.121	-.161	-.154	.79	.28	.02	.12	.05	.15	.788	.185	.054	-.013	.278	-.051
Pk. up cup, secure cube	84	-.886	-.019	.070	.147	.172	-.029	.76	.00	-.09	.47	.12	.09	.921	-.082	-.007	.042	-.058	.059
Cube in cup on command	88	.575	.114	-.081	.450	.066	-.123	.48	.14	-.12	.29	.20	.43	.527	.019	-.197	.413	.050	-.055
Unwraps cube	90	.761	.135	.178	.038	.122	-.148	.71	.13	.06	.31	.18	-.02	.771	.084	.097	-.096	-.030	-.061
Holds crayon adaptively	91	.525	-.008	.636	.084	.147	-.131	.56	-.09	.53	.28	.17	-.04	.531	-.143	.611	-.014	-.072	-.012
Inhib. on command	92	.046	.140	.655	.045	.132	-.056	.13	.04	.64	.12	.10	-.07	-.000	.048	.682	-.015	-.100	.042
Strikes doll in imit.	94	.218	.173	.132	.218	.193	-.803	.30	.16	.15	-.11	.78	.15	.097	.093	-.066	.148	-.182	-.803
Imitates words	95	.500	.526	-.050	.425	.381	.224	.24	.54	.00	.68	.04	.31	.371	.537	-.090	.330	-.231	.352
Stirs with spn. in imit.	96	.426	.219	.285	.516	-.174	-.304	.51	.18	.31	.00	.22	.51	.349	.023	.158	.476	.290	-.203
Holds cup to drink from	97	.100	.113	.318	.188	.643	-.253	-.02	.06	.32	.44	.54	-.04	-.025	.101	.269	.132	-.619	-.201
Says two words	99	.273	.809	.050	-.024	.159	-.192	.24	.80	.14	.12	.22	-.07	.132	.882	-.028	-.202	-.092	-.113
Dangles ring by string	100	.110	.052	.083	.663	.171	-.124	.04	.03	.15	.26	.26	.58	-.008	-.086	-.000	.712	-.083	-.062
Uses expressive jargon	102	-.083	.715	.467	.379	-.134	-.115	.04	.61	.61	-.08	.07	.33	-.256	.604	.428	.309	.225	.016
LAMBDA		3.164	1.664	1.330	1.363	.811	.998												
PERCENT		24.34	12.80	10.23	10.48	6.24	7.68												

TABLE 11. FACTOR ANALYSIS, MERRILL-PALMER SCALE
AGE LEVEL, 24 MONTHS

Item Name	Item No.	Test Item Intercorrelations												Principal Axes (Unrotated)						h^2
		7	10	11	13	14	17	18	19	21	23	26	28	1	2	3	4	5	6	
Cross. feet	7		745	480	517	136	219	188	394	287	101	228	046	-.693	-.391	-.423	-.135	-.036	-.074	.766
Stand on one ft.	10	745		454	443	154	289	357	408	301	227	245	168	-.698	-.221	-.454	-.029	-.045	-.248	.807
Folding paper	11	480	454		380	111	298	119	541	520	025	192	040	-.600	-.371	-.032	.123	-.041	.293	.601
Ident. in mirror	13	517	443	380		130	203	159	662	473	325	454	230	-.691	-.103	.071	-.272	.178	.056	.601
16 cubes	14	136	154	111	130		121	322	073	-105	328	258	037	-.234	.425	-.338	-.162	-.259	.325	.549
Drawing up string	17	219	289	298	203	121		387	025	306	163	173	112	-.375	.121	-.267	.483	.309	.157	.581
Nest of cubes	18	188	357	119	159	322	387		380	389	397	259	239	-.557	.481	-.115	.358	-.216	-.208	.773
Questions	19	394	408	541	662	073	025	380		772	403	393	240	-.832	-.097	.451	-.152	-.213	-.029	.975
Rep. of word gr.	21	287	301	520	473	-105	306	389	772		358	245	131	-.735	-.131	.473	.340	-.006	.029	.897
Match colors	23	101	227	025	325	328	163	397	403	358		445	224	-.484	.475	.095	-.118	-.009	-.006	.483
Wallin B	26	228	245	192	454	258	173	259	393	245	445		335	-.511	.320	.021	-.288	.252	.075	.516
Wallin A	28	046	168	040	230	037	112	239	240	131	224	335		-.296	.239	.064	-.106	.240	-.192	.255
													LAMBDA	4.053	1.188	1.031	.743	.413	.374	
													PERCENT	33.77	9.90	8.59	6.19	3.44	3.11	

Item Name	Item No.	Varimax Rotation						Graphic Rotation (Orthogonal)						Biquartimin Rotation					
		Factors						Factors						Factors					
		1	2	3	4	5	6	A	B	C	D	E	F	1	2	3	4	5	6
Cross feet	7	-.838	.084	.207	.084	.019	.080	78	02	17	28	04	20	.839	.000	.066	-.014	-.050	-.029
Stand on one ft.	10	-.829	.167	.142	.131	-.229	.051	78	23	18	27	17	03	.848	-.074	.004	-.059	.188	-.002
Folding paper	11	-.429	-.045	.568	.259	.117	.108	32	-21	44	43	06	23	.332	.157	.522	-.190	-.168	-.061
Ident. in mirror	13	-.403	.484	.433	.034	.111	.067	37	20	48	09	-07	45	.307	-.442	.325	.033	-.233	.002
16 cubes	14	-.089	.127	-.051	.053	-.123	.710	04	15	-11	07	66	32	.041	-.057	-.110	-.017	.081	-.722
Drawing up string	17	-.154	.106	.084	.714	-.158	.060	00	27	12	68	08	00	.085	-.077	.011	-.711	.069	-.018
Nest of cubes	18	-.130	.223	.179	.253	-.744	.237	10	48	33	21	48	-23	.102	-.098	.177	-.196	.681	-.229
Questions	19	-.267	.362	.833	-.204	-.185	.046	32	07	89	-15	14	21	.151	-.222	.843	.315	.094	.003
Rep. of word gr.	21	-.098	.185	.842	.232	-.249	-.169	08	05	91	27	04	-06	-.034	-.058	.887	-.153	.155	.222
Match colors	23	.007	.513	.207	.025	-.296	.297	00	46	32	-06	31	18	-.086	-.465	.167	.023	.183	-.272
Wallin B	26	-.124	.643	.161	.086	.004	.234	08	47	27	02	09	46	.029	-.647	.051	-.050	-.150	-.182
Wallin A	28	-.055	.474	.028	.050	-.142	-.059	05	42	16	-04	-07	10	.027	-.498	-.033	-.037	.051	.094
LAMBDA		1.883	1.414	2.088	.775	.877	.766												
PERCENT		15.69	11.78	17.40	6.46	7.31	6.38												

TABLE 10. FACTOR ANALYSIS, GESELL DEVELOPMENTAL SCHEDULE AGE LEVEL, 12 MONTHS

Item Name	Item No.	Test Item Intercorrelations																			Principal Axes (Unrotated)								
		M15 (A)	M20 (B‡)	L11 (B)	A23 (‡)	M35 (B)	M34 (B_1)	P41 (A)	A40 (C)	M40 (B)	A32 (B‡)	A52 (B)	A22 (A)	A50 (B)	M34 (B_2)	P14 (B‡)	P31 (A)	P22	P20	P21	1	2	3	4	5	6	7	8	h^2
Stands alone	M15(A)		500	000	492	192	203	459	378	287	459	278	327	391	203	012	315	087	340	409	537	248	-067	-333	-006	-108	-200	-165	.544
Can climb	M20(B‡)	500		137	113	153	343	420	554	541	356	143	482	291	416	004	262	043	495	541	602	324	-491	-102	045	-277	070	131	.820
Says 4 words	L11(B)	000	137		288	-064	274	000	067	061	000	183	063	-155	039	661	390	363	258	211	267	-888	-220	-032	-034	-293	092	-140	1.000
Builds tower of two	A23(‡)	492	113	288		359	294	552	358	382	382	398	008	300	294	300	416	223	399	422	604	-183	184	-390	-225	186	-232	-346	.843
Takes 3rd cube	M35(B)	192	153	-064	359		567	347	421	222	409	612	277	347	117	096	173	126	094	383	511	133	364	135	399	237	012	-180	.677
Inhibits hand to mouth	M34(B_1)	203	343	274	294	567		698	511	495	495	401	319	434	440	521	300	494	538	530	791	-174	049	394	050	219	-039	284	.947
Puts cube in cup-play	P41(A)	459	420	000	552	347	698		641	565	565	259	403	361	495	420	292	346	545	449	813	116	-080	208	-320	206	-355	-028	.996
Puts cube in cup-com.	A40(C)	378	554	067	358	421	511	641		641	413	400	362	270	298	200	108	252	682	395	723	205	-218	151	-128	204	302	-212	.829
Scribble imitation	M40(B)	287	541	061	382	222	495	565	641		565	434	345	534	349	178	352	287	545	219	713	201	019	048	-290	-151	291	068	.747
Unwraps cube	A32(B‡)	459	356	000	382	409	495	565	413	565		543	345	534	349	178	352	227	232	098	650	225	322	076	-126	-228	-031	-071	.656
Shown form-board circle	A52(B)	278	143	183	398	612	401	259	400	434	543		290	584	237	366	426	339	393	321	650	-096	458	-039	269	-098	297	-159	.838
Hits rubber doll	A22(A)	327	482	063	008	277	319	403	362	345	345	290		089	244	283	075	-050	160	418	473	215	-288	349	376	-359	-182	-237	.834
Performance box	A50(B)	391	291	-155	300	347	434	361	270	534	534	584	089		348	265	448	204	217	350	594	167	428	-151	023	-127	044	300	.696
Preference for one hand	M34(B_2)	203	416	039	294	117	440	495	298	349	349	237	244	348		262	300	125	162	215	481	071	006	086	-161	-167	-197	177	.367
Says "bye bye"	P14(B‡)	012	004	661	300	096	521	420	200	178	178	366	283	265	262		151	330	280	231	445	-572	035	309	-057	-107	-089	-043	.647
Asks for things at table	P31(A)	315	262	390	416	173	300	292	108	352	352	426	075	448	300	151		191	295	462	507	-197	140	-405	029	-220	-047	171	.560
Bowel control	P22	087	043	363	223	126	494	346	252	287	227	339	-050	204	125	330	191		072	409	388	-334	111	082	-022	121	003	173	.326
Tries to put on shoes	P20	340	495	258	399	094	538	545	682	545	232	393	160	217	162	280	295	072		664	665	-065	-388	-164	-108	239	311	-026	.790
Uses spoon	P21	409	541	211	422	383	530	449	395	219	098	321	418	350	215	231	462	409	664		683	-130	-346	-300	469	264	-137	192	1.000
																				LAMBDA	6.833	1.770	1.400	1.058	.910	.855	.699	.654	
																				PERCENT	35.96	9.32	7.37	5.57	4.79	4.50	3.68	3.44	

Item Name	Item No.	Varimax Rotation								Graphic Rotation (Orthogonal)						Biquartimin Rotation							
		Factors								Factors						Factors							
		1	2	3	4	5	6	7	8	A	B	C	D	E	F	1	2	3	4	5	6	7	8
Stands alone	M15(A)	.211	-.099	-.052	-.496	-.208	.307	.017	-.323	.27	-.03	.18	.61	.12	.14	-.094	.170	-.018	.460	.091	-.238	.032	-.234
Can climb	M20(B‡)	.512	-.027	.163	-.032	-.328	.526	.046	-.343	.08	-.08	.48	.62	.10	.40	-.479	.075	-.294	-.094	.188	-.442	.031	-.281
Says 4 words	L11(B)	.077	.984	.099	-.094	-.112	.026	-.071	.007	-.12	.91	.06	.08	-.26	.23	-.047	-1.015	-.108	.050	.016	-.066	.200	.046
Builds tower of two	A23(‡)	.179	.227	-.209	-.799	-.119	-.115	.163	-.195	.47	.39	.03	.45	-.05	-.29	-.079	-.139	.142	.805	.000	.207	-.103	-.016
Takes 3rd cube	M35(B)	.077	-.067	-.753	-.129	-.151	.138	.195	-.101	.38	.06	-.01	-.02	.72	.06	.008	.144	.796	.081	.144	-.140	-.008	.050
Inhibits hand to mouth	M34(B_1)	.299	.299	-.350	.080	-.262	.084	.722	-.228	.31	.46	.52	.01	.56	.02	-.225	-.238	.156	-.219	.216	.019	-.644	-.096
Puts cube in cup-play	P41(A)	.383	.027	-.077	-.434	-.073	.242	.750	-.175	.41	.18	.70	.26	.29	-.16	-.283	.012	-.188	.370	-.002	-.181	-.794	.065
Puts cube in cup-com	A40(C)	.789	.005	-.275	-.161	-.051	.215	.245	-.060	.24	.07	.62	.34	.38	-.05	-.884	.051	.180	.065	-.070	-.091	-.069	.245
Scribble imitation	M40(B)	.615	.066	-.109	-.081	.084	.143	.226	-.511	.54	.07	.52	.29	.08	.08	-.640	.001	-.034	-.040	-.271	.003	-.060	-.365
Unwraps cube	A32(B‡)	.211	.007	-.311	-.243	.176	.253	.261	-.540	.71	.03	.29	.13	.16	.19	-.097	.044	.209	.175	-.313	-.168	-.117	-.424
Shown form-board circle	A52(B)	.207	.243	-.711	-.117	-.033	.062	-.030	-.465	.62	.32	.07	.10	.43	.24	-.137	-.144	.771	.026	-.102	-.047	.355	-.356
Hits rubber doll	A22(A)	.135	.064	-.204	.001	-.090	.856	.127	-.038	.14	-.02	.41	.10	.29	.66	-.003	-.106	.163	-.064	.059	-.900	-.013	.105
Performance box	A50(B)	.102	-.067	-.299	-.077	-.119	-.009	.206	-.724	.70	.06	.01	.22	.24	.12	.020	.186	.207	-.029	-.010	.105	-.085	-.769
Preference for one hand	M34(B_2)	.117	.068	.042	-.120	-.032	.217	.373	-.374	.35	.11	.36	.15	.04	.14	-.006	-.044	-.210	.052	-.049	-.132	-.382	-.345
Says "bye bye"	P14(B‡)	.038	.689	-.145	-.054	.023	.104	.389	-.061	.15	.70	.29	-.15	.07	.14	.047	-.716	.045	-.001	-.072	-.097	-.306	.044
Asks for things at table	P31(A)	.043	.278	-.049	-.259	-.301	-.011	-.005	-.556	.41	.37	-.11	.43	.02	.17	.090	-.180	-.013	.179	.159	.066	.097	-.624
Bowel control	P22	.084	.337	-.166	-.013	-.163	-.147	.327	-.165	.19	.44	.11	.00	.19	-.08	-.041	-.291	.079	-.054	.124	.198	-.279	-.136
Tries to put on shoes	P20	.749	.184	-.050	-.183	-.360	.009	.118	-.086	.04	.30	.42	.58	.26	-.14	-.848	-.093	-.047	.071	.223	.119	.012	.109
Uses spoon	P21	.241	.136	-.189	-.185	-.904	.189	.178	-.117	.06	.37	-.03	.56	.65	.09	-.132	-.007	.114	.061	.859	-.142	-.128	-.088
	LAMBDA	2.345	1.920	1.690	1.368	1.403	1.454	1.836	2.151														
	PERCENT	12.34	10.10	8.90	7.20	7.39	7.65	9.66	11.32														

TABLE 9. FACTOR ANALYSIS, GESELL DEVELOPMENTAL SCHEDULE
AGE LEVEL, 6 MONTHS

Item Name	Item No.	Test Item Intercorrelations															
		M 22	M 20	M 25	P41 (B_1)	A17 (B)	P41 (B_2)	A 16	M34 (A_2)	A11 (B)	A13 (B+)	M32 (B)	A31 (B)	A30 (A)	M35 (B)	A15 (B)	A18 (B)
Rolls back to stomach	M 22		465	001	206	144	220	397	-009	050	-274	441	144	-287	-050	179	232
Creeps or hitches	M 20	465		392	158	313	422	294	437	029	121	314	189	143	-305	248	369
Makes stepping movement	M 25	001	392		129	398	153	356	179	293	257	564	398	259	142	400	232
Pats table	$P41(B_1)$	206	158	129		335	725	453	114	291	083	406	335	116	083	363	273
Manipulates spoon	A17(B)	144	313	398	335		478	492	583	-003	281	799	719	138	255	625	652
Bangs spoon	$P41(B_2)$	220	422	153	725	478		404	368	-067	251	491	548	-039	241	152	194
Reaches directly	A16	397	294	356	453	492	404		699	017	149	730	492	015	217	247	720
Manipul. spn., 1 hand	$M34(A_2)$	-009	437	179	114	583	368	699		032	322	391	248	088	430	251	484
Blinks at pencil	A11(B)	050	029	293	291	-003	-067	017	032		320	066	-174	-030	-091	116	-231
Regards pellet	A13(B+)	-274	121	257	083	281	251	149	322	320		273	281	484	368	353	273
Picks up cube	M32(B)	441	314	564	406	799	491	730	391	066	273		754	243	194	344	697
Lifts inverted cup	A31(B)	144	189	398	335	719	548	492	248	-174	281	754		388	338	496	535
Looks for fallen obj.	A30(A)	-287	143	259	116	138	-039	015	088	-030	484	243	388		026	218	253
Holds 2 cubes	M35(B)	-050	-305	142	083	255	241	217	430	-091	368	194	338	026		484	288
Dangling ring	A15(B)	179	248	400	363	625	152	247	251	116	353	344	496	218	484		430
Crumples paper	A18(B)	232	369	232	273	652	194	720	484	-231	273	697	535	253	288	430	
Purposeful reaction	A18(A)	283	073	315	024	098	185	661	491	153	009	196	098	-279	167	055	471
Expresses recognition	P 11	159	280	361	127	482	321	260	432	084	608	359	217	538	417	402	422
Splashes in tub	M 23	-126	272	179	-376	017	-141	307	411	172	371	096	017	374	305	141	314

Item Name	Item No.	Principal Axes (Unrotated)								h^2
		1	2	3	4	5	6	7	8	
Rolls back to stomach	M 22	.296	-.616	.129	.217	.092	.170	.355	.380	.838
Creeps or hitches	M 20	.483	-.195	.118	.721	.265	-.325	.200	-.097	1.000
Makes stepping movement	M 25	.517	.120	.003	.251	-.093	.298	.015	-.204	.484
Pats table	$P41(B_1)$	.466	-.418	-.519	.019	-.428	-.126	-.243	.194	.957
Manipulates spoon	A17(B)	.813	-.042	-.214	-.121	.125	.065	.157	-.332	.878
Bangs spoon	$P41(B_2)$	.582	-.338	-.320	-.054	-.176	-.469	-.045	-.034	.812
Reaches directly	A16	.795	-.306	.381	-.101	-.039	.072	-.345	.097	1.000
Manipul. spn., 1 hand	$M34(A_2)$	.681	.105	.444	-.094	-.079	-.396	-.034	-.201	.885
Blinks at pencil	A11(B)	.075	.089	.031	.429	-.719	.280	-.013	-.102	.805
Regards pellet	A13(B+)	.450	.581	-.100	.090	-.248	-.106	.008	.019	.631
Picks up cube	M32(B)	.837	-.188	-.132	.020	.141	.292	-.108	-.076	.877
Lifts inverted cup	A31(B)	.728	-.020	-.340	-.210	.245	.160	-.083	-.128	.798
Looks for fallen obj.	A30(A)	.319	.676	-.341	.282	.264	.013	-.317	.278	1.000
Holds 2 cubes	M35(B)	.419	.337	.070	-.589	-.225	-.060	.301	.138	.805
Dangling ring	A15 (B)	.594	.161	-.202	-.057	-.082	.183	.377	.007	.606
Crumples paper	A18(B)	.766	-.011	.186	-.127	.293	.094	-.087	.146	.760
Purposeful reaction	A18(A)	.399	-.208	.629	-.098	-.215	.091	-.109	.069	.679
Expresses recognition	P 11	.598	.368	-.038	.104	-.075	-.136	.220	.258	.644
Splashes in tub	M 23	.276	.487	.509	.149	.042	.003	-.060	.022	.600
LAMBDA		6.148	2.220	1.785	1.392	1.250	.899	.795	.616	
PERCENT		32.36	11.69	9.39	7.32	6.58	4.73	4.18	3.24	

Item Name	Item No.	Varimax Rotation								Graphic Rotation (Orthogonal)							
		Factors								Factors							
		1	2	3	4	5	6	7	8	A	B	C	D	E	F	G	H
Rolls back to stomach	M 22	.141	-.193	.169	.221	.004	-.134	-.026	.828	.37	-.21	.06	-.22	.09	.28	-.29	.65
Creeps or hitches	M 20	.202	.135	.133	.937	-.057	-.083	-.103	.232	.08	.07	.39	.42	.22	.07	.26	.76
Makes stepping movement	M 25	.496	.146	.156	.166	-.396	.058	.070	.011	.42	.22	.28	.12	.35	.00	.21	.06
Pats table	$P41(B_1)$	.228	.076	.052	-.070	-.208	-.908	.006	.154	.26	.13	-.26	.07	.20	.87	.04	-.03
Manipulates spoon	A17(B)	.826	-.035	.119	.230	.020	-.196	.293	-.059	.59	.42	.19	.38	-.06	.22	.10	.21
Bangs spoon	$P41(B_2)$	.264	-.054	.139	.294	.087	-.767	.191	-.038	.10	.22	-.02	.40	-.02	.75	-.02	.19
Reaches directly	A16	.427	.054	.856	.039	-.009	-.263	.000	.163	.43	.05	.71	-.06	-.06	.54	-.07	-.01
Manipul. spn., 1 hand	$M34(A_2)$	.205	-.010	.665	.408	.044	-.101	.382	-.275	-.07	.40	.71	.32	-.07	.39	-.02	.14
Blinks at pencil	A11(B)	-.064	.016	.048	.008	-.890	-.080	.023	-.002	-.01	.24	.05	-.19	.83	.08	.04	-.11
Regards pellet	A13(B+)	.155	.459	.073	.119	-.280	-.061	.476	-.261	-.06	.53	.19	.06	.15	.09	.51	.01
Picks up cube	M32(B)	.813	.143	.313	.084	-.064	-.223	.018	.191	.77	.14	.31	.18	.03	.32	.15	.09
Lifts inverted cup	A31(B)	.808	.173	.095	-.020	.161	-.244	.142	-.015	.67	.27	.10	.30	-.23	.24	.23	.02
Looks for fallen obj.	A30(A)	.212	.957	-.115	150	.014	.025	.065	-.142	.09	.04	.06	.00	-.07	.02	.98	-.03
Holds 2 cubes	M35(B)	.160	-.008	.207	242	.091	-.030	.814	-.077	.04	.80	.14	-.18	-.33	.09	-.02	.06
Dangling ring	A15 (B)	.518	.095	-.041	.076	-.169	-.059	.511	.168	.42	.54	.01	.03	.07	.07	.16	.30
Crumples paper	A18(B)	.546	.248	.523	.096	.215	-.028	.186	.191	.49	.15	.53	.01	-.29	.25	.17	.19
Purposeful reaction	A18(A)	.051	-.170	.770	.001	-.149	.032	.094	.152	.10	.11	.66	-.24	.06	.23	-.31	.01
Expresses recognition	P 11	.199	.425	.139	.250	-.095	-.093	.559	.110	.06	.46	.22	-.04	.00	.20	.44	.40
Splashes in tub	M 23	-.011	.337	.444	.195	-.125	.411	.215	-.145	-.12	.18	.64	-.09	.05	-.20	.30	.06
LAMBDA		3.355	1.660	2.536	1.465	1.243	1.862	1.862	1.121								
PERCENT		17.66	8.73	13.39	7.71	6.54	9.80	9.80	5.90								

TABLE 12. FACTOR ANALYSIS, MERRILL-PALMER SCALE
AGE LEVEL, 30 MONTHS

Item Name	Item No.	Test Item Intercorrelations													Principal Axes (Unrotated)					
		20	22	23	25	29	32	34	35	36	37	38	40	43	1	2	3	4	5	h²
Scissors	20		230	081	-013	093	361	595	276	237	179	257	172	171	.464	-.030	-.394	-.345	-.424	.671
One button	22	230		357	088	349	025	445	-074	118	228	143	048	093	.327	.227	-.248	.284	-.172	.330
Match colors	23	081	357		233	359	291	411	311	440	464	321	179	147	.572	.136	-.102	.176	.358	.515
16 cubes	25	-013	088	233		274	247	365	202	227	464	236	299	533	.510	-.180	.235	.251	.032	.412
Nest cubes	29	093	349	359	274		147	350	-223	394	353	363	384	002	.511	.532	.104	.338	-.082	.676
Word groups	32	361	025	291	247	147		324	337	346	242	471	308	336	.529	-.057	.078	-.286	.017	.371
Fist, thumb	34	595	445	411	365	350	324		525	334	554	334	205	134	.803	-.020	-.582	.043	-.079	.991
Two blocks	35	276	-074	311	202	-223	337	525		180	308	004	298	297	.446	-.562	-.242	-.292	.325	.764
Seguin (time)	36	237	118	440	227	394	346	334	180		221	629	350	131	.602	.338	.126	-.214	.225	.588
Circle	37	179	228	464	464	353	242	554	308	221		090	162	538	.634	-.245	-.069	.457	.058	.678
Wallin A	38	257	143	321	236	363	471	334	004	629	090		507	118	.580	.443	.250	-.354	.013	.720
Little pink tow.	40	172	048	179	299	384	308	205	298	350	162	507		566	.582	-.026	.444	-.185	-.121	.586
Three cube pyr.	43	171	093	147	533	002	336	134	297	131	538	118	566		.522	-.557	.462	.138	-.248	.877
												LAMBDA			4.004	1.387	1.191	1.010	.587	
												PERCENT			30.80	10.67	9.16	7.77	4.52	

Item Name	Item No.	Varimax Rotation					Graphic Rotation (Orthogonal)					Biquartimin Rotation				
		Factors					Factors					Factors				
		1	2	3	4	5	A	B	C	D	E	1	2	3	4	5
Scissors	20	.047	-.119	-.785	-.181	.075	44	00	53	-09	46	-.016	-.097	-.750	.137	-.130
One button	22	.005	.174	-.251	-.026	.486	19	-10	21	48	18	-.022	.191	-.180	-.047	-.566
Match colors	23	.063	-.250	.060	-.331	.379	36	15	26	41	-35	-.045	-.249	.187	.290	-.462
16 cubes	25	.521	-.095	.051	-.162	.320	10	52	10	34	-05	.492	-.033	.133	.088	-.258
Nest cubes	29	.106	.388	-.027	-.408	.589	47	05	-13	66	04	.069	.426	.072	.362	-.565
Word groups	32	.241	-.247	-.233	-.441	.058	44	35	21	-07	00	.162	-.217	-.175	.438	.069
Fist, thumb	34	.082	-.335	-.631	-.178	.665	45	08	76	36	08	-.050	-.320	-.490	.067	-.688
Two blocks	35	.208	-.819	-.199	-.082	.061	11	42	71	-28	-33	.093	-.824	-.131	.051	.062
Seguin (time)	36	.036	-.113	-.044	-.705	.274	68	19	01	14	-21	-.075	-.104	.049	.734	-.079
Circle	37	.484	-.209	-.069	.015	.629	05	45	41	57	-06	.429	-.154	.057	-.147	-.629
Wallin A	38	.067	.076	-.156	-.821	.105	84	21	-14	05	-01	-.020	.101	-.092	.871	.083
Little pink tow.	40	.524	-.007	-.111	-.547	-.002	45	59	-07	04	09	.487	.067	-.061	.540	.147
Three cube pyr.	43	.923	-.121	-.081	-.064	.012	-05	92	10	13	26	.941	-.008	-.039	-.030	.028
LAMBDA		1.762	1.188	1.227	2.044	1.958										
PERCENT		13.56	9.14	9.44	15.72	15.06										

TABLE 13. FACTOR ANALYSIS, MERRILL-PALMER SCALE
AGE LEVEL, 36 MONTHS

Item Name	Item No.	Test Item Intercorrelations 25	32	34	36	37	39	41	42	44	45	48	68	Principal Axes (Unrotated) 1	2	3	4	5	h^2
16 cubes	25		574	642	651	414	168	576	599	619	519	444	234	.849	-.151	-.201	-.105	-.156	.823
Word groups	32	574		421	559	502	403	503	301	508	673	418	233	.787	.205	.036	.369	.049	.801
Fist, thumb	34	642	421		390	445	038	477	412	649	233	314	-003	.695	-.564	-.187	-.051	.356	.966
Seguin (time)	36	651	559	390		429	400	735	409	504	424	234	278	.776	[illegible]	.223	-.248	-.023	.731
Circle	37	414	502	445	429		338	435	249	327	108	436	101	.574	-.205	.373	.257	-.042	.579
Pic. Puz. #1	39	168	403	038	400	338		395	103	253	162	032	199	.375	.245	.448	-.009	.021	.401
Seguin (errors)	41	576	503	477	735	435	395		302	411	463	196	310	.743	.156	.226	-.248	.179	.721
Two buttons	42	599	301	412	409	249	103	302		386	210	230	251	.530	-.173	-.151	-.242	-.327	.500
Mare and foal	44	619	508	649	504	327	253	411	386		322	363	241	.681	-.182	-.073	-.023	.011	.503
Wallin and Peg B	45	519	673	233	424	108	162	463	210	322		171	415	.632	.597	-.450	.136	.138	.996
Three cube pyr.	48	444	418	314	234	436	032	196	230	363	171		101	.454	-.227	-.024	.365	-.239	.449
Little pink tow.	68	234	233	-003	278	101	199	310	251	241	415	101		.336	.325	-.056	-.119	-.192	.273
													LAMBDA	4.911	1.103	.756	.564	.409	
													PERCENT	40.93	9.19	6.30	4.70	3.41	

Item Name	Item No.	Varimax Rotation Factors 1	2	3	4	5	Graphic Rotation (Orthogonal) Factors A	B	C	D	E	Biquartimin Rotation Factors 1	2	3	4	5
16 cubes	25	.464	.310	.203	.343	-.594	.40	.50	.13	.50	.43	-.313	-.186	-.056	.203	.533
Word groups	32	.227	.538	.407	.538	-.075	.56	.51	.46	.12	.02	-.093	-.514	-.295	.446	-.111
Fist, thumb	34	.920	.041	.078	.271	-.195	.54	.17	-.09	.81	.04	-.922	-.015	.010	.058	.099
Seguin (time)	36	.273	.240	.641	.132	-.414	.11	.36	.56	.48	.27	-.152	-.067	-.610	-.061	.330
Circle	37	.236	-.051	.441	.566	-.070	.56	-.05	.45	.24	.19	-.141	.143	-.435	.472	-.028
Pic. puz. #1	39	-.047	.092	.611	.131	-.016	.05	.05	.62	.12	.06	.113	.011	-.660	.033	-.066
Seguin (errors)	41	.367	.273	.666	.041	-.260	.09	.34	.57	.54	.08	-.292	-.137	-.652	-.181	.147
Two buttons	42	.226	.074	.088	.180	-.635	.14	.27	.03	.36	.58	-.094	.072	.009	.104	.665
Mare and foal	44	.457	.191	.224	.308	-.336	.39	.25	.23	.46	.23	-.367	-.110	-.132	.171	.259
Wallin and Peg B	45	.140	.964	.145	.076	-.145	.13	.95	.18	.03	-.12	-.037	-1.014	.035	.009	-.026
Three cube pyr.	48	.146	.085	.018	.619	-.191	.59	.12	.08	.04	.30	-.024	-.051	.079	.628	.135
Little pink tow.	68	-.105	.338	.221	.009	-.315	-.09	.40	.24	.02	.24	.210	-.260	-.161	-.022	.290
	LAMBDA	1.692	1.622	1.762	1.354	1.313										
	PERCENT	14.10	13.51	14.69	11.29	10.94										

TABLE 14. FACTOR ANALYSIS, MERRILL-PALMER SCALE AGE LEVEL, 48 MONTHS

Item Name	Item No.	Test Item Intercorrelations																Principal Axes (Unrotated)						
		50	54	55	56	59	62	63	68	71	72	73	76	77	78	80	81	1	2	3	4	5	6	h^2
Seguin errors	50		424	416	529	221	058	212	260	411	656	427	243	624	596	331	404	.740	.145	.090	-.056	-.163	-.285	.683
Thumb, finger	54	424		367	438	426	-052	322	100	129	231	-039	154	243	253	270	128	.421	-.268	.382	.014	-.007	-.228	.448
4 buttons	55	416	367		316	271	040	520	173	396	490	144	216	400	334	-047	185	.539	-.073	.448	-.323	-.263	.214	.716
Cross	56	529	438	316		445	273	318	169	399	447	302	052	487	636	382	301	.698	.044	.260	.068	.372	-.274	.775
Pic. puz. #1	59	221	426	271	445		158	281	-029	432	328	207	646	528	218	322	186	.604	-.692	-.069	-.055	.317	.164	.979
3 cube pyr.	62	058	-052	040	273	158		105	174	052	169	326	-128	316	166	225	178	.257	.216	-.043	.252	.362	.226	.360
2 buttons	63	212	322	520	318	281	105		252	103	207	169	-038	135	227	195	-006	.363	-.062	.590	.084	.002	.339	.606
Little pink twr.	68	260	100	173	169	-029	174	252		-014	321	226	146	350	339	490	228	.392	.218	.052	.540	-.355	.143	.642
Pic. puz. #3	71	411	129	396	399	432	052	103	-014		602	508	365	453	506	243	528	.668	.012	-.251	-.378	.054	.062	.658
Decroly	72	656	231	490	447	328	169	207	321	602		510	135	531	504	253	677	.771	.286	-.054	-.221	-.093	.071	.741
Manikin	73	427	-039	144	302	207	326	169	226	508	510		101	320	361	283	490	.542	.326	-.242	-.035	.145	.245	.540
Pic. puz. #2	76	243	154	216	052	646	-128	-038	146	365	135	101		434	206	361	210	.431	-.627	-.396	.013	-.312	.028	.834
Seguin time	77	624	243	400	487	528	316	135	350	453	531	320	434		579	551	409	.778	-.079	-.127	.177	-.029	-.067	.664
Mare & foal	78	596	253	334	636	218	166	227	339	506	504	361	206	579		236	417	.697	.188	.053	-.014	-.014	-.209	.568
6 cube pyr.	80	331	270	-047	382	322	225	195	490	243	253	283	361	551	236		193	.517	-.113	-.148	.579	.009	-.029	.639
Action agent	81	404	128	185	301	186	178	-006	228	528	677	490	210	409	417	193		.585	.289	-.312	-.176	-.037	.033	.551
																	LAMBDA	5.438	1.395	1.199	1.070	.725	.588	
																	PERCENT	33.99	8.72	7.50	6.69	4.53	3.67	

Item Name	Item No.	Varimax Rotation						Graphic Rotation (Orthogonal)						Biquartimin Rotation					
		Factors						Factors						Factors					
		1	2	3	4	5	6	A	B	C	D	E	F	1	2	3	4	5	6
Seguin errors	50	.552	-.055	.152	.257	-.123	-.525	.29	.25	.16	.11	.51	-.22	.481	.065	-.034	.161	-.237	-.490
Thumb, finger	54	.002	-.200	.313	.061	-.096	-.545	-.14	.50	.14	.18	.25	-.20	-.106	-.126	-.231	.009	-.152	-.531
4 buttons	55	.347	-.110	.719	.005	-.179	-.186	.07	.42	.09	.08	.71	.20	.292	-.048	-.701	-.029	-.214	-.043
Cross	56	.305	-.083	.159	.052	.316	-.740	.31	.63	.08	.48	.00	-.21	.214	.057	-.028	-.077	.211	-.758
Pic. puz. #1	59	.129	-.858	.235	-.120	.263	-.295	.01	.64	.58	-.02	-.06	.19	.033	-.833	-.177	-.211	.285	-.164
3 cube pyr.	62	.152	.025	.021	.146	.559	-.048	.25	.08	-.03	.37	-.20	.17	.136	.046	-.011	.102	.538	-.011
2 buttons	63	-.017	-.021	.714	.147	.199	-.185	-.16	.41	-.11	.38	.36	.34	-.110	.007	-.721	.139	.178	-.065
Little pink twr.	68	.152	.043	.171	.761	.080	-.037	-.08	-.28	.20	.70	.26	.04	.152	.037	-.154	.777	.018	.068
Pic. puz. #3	71	.714	-.322	.094	-.111	.053	-.147	.59	.27	.40	-.01	.21	.09	.727	-.242	-.030	-.219	.014	-.048
Decroly	72	.778	-.055	.244	.150	.061	-.216	.61	.18	.23	.27	.45	.05	.765	.046	-.174	.052	-.022	-.108
Manikin	73	.618	-.047	.045	.140	.367	.013	.60	.03	.17	.29	.03	.30	.634	.005	-.017	.062	.324	.108
Pic. puz. #2	76	.191	-.822	-.029	.218	-.270	.000	-.09	.08	.90	-.11	.12	.01	.114	-.855	.067	.193	-.246	.142
Seguin time	77	.459	-.379	.073	.376	.129	-.382	.28	.26	.53	.43	.17	-.09	.561	-.309	.026	.284	.054	-.287
Mare & foal	78	.527	-.034	.112	.217	.045	-.477	.42	.27	.19	.39	.31	-.29	.469	.080	-.005	.118	-.060	-.448
6 cube pyr.	80	.098	-.339	-.073	.599	.272	-.277	.00	.05	.48	.63	-.14	-.01	-.027	-.326	.143	.556	.219	-.207
Action agent	81	.724	-.077	-.042	.117	.085	-.070	.64	-.02	.27	.15	.21	-.02	.746	-.006	.098	.031	.025	.003
LAMBDA		3.137	1.852	1.353	1.373	.897	1.803												
PERCENT		19.60	11.57	8.46	8.58	5.61	11.27												

TABLE 15. FACTOR ANALYSIS, MERRILL-PALMER SCALE
AGE LEVEL, 54 MONTHS

Item Name	Item No.	Test Item Intercorrelations												Principal Axes (Unrotated)					h^2
		55	69	75	76	78	80	82	85	86	88	89	92	1	2	3	4	5	
Four buttons	55		840	327	428	244	184	196	166	399	184	059	414	.590	.607	.341	.229	.055	.889
Three cube pyr.	69	840		508	379	360	141	071	318	266	396	053	443	.656	.653	.203	-.127	.031	.915
Two buttons	75	327	508		235	345	043	117	400	141	377	213	345	.495	.235	-.035	-.449	-.093	.511
Pict. puz. #2	76	428	379	235		580	297	420	651	297	569	160	463	.751	-.016	-.314	.208	.355	.833
Mare and foal	78	244	360	345	580		359	480	473	449	478	613	446	.757	-.315	.088	-.162	.216	.752
Six cube pyr.	80	184	141	043	297	359		213	149	474	182	340	179	.385	-.244	.260	.193	.081	.320
Copy star	82	196	071	117	420	480	213		439	390	247	171	390	.494	-.229	-.121	.225	-.057	.365
Action agent	85	166	318	400	651	473	149	439		338	576	149	591	.722	-.098	-.510	-.060	-.164	.821
Sequin form bd.	86	339	266	141	297	449	474	390	338		361	474	493	.621	-.269	.310	.226	-.265	.675
Decroly matching	88	184	396	377	569	478	182	247	576	361		136	376	.618	-.014	-.260	-.111	.050	.465
Little pink tow.	89	059	053	213	160	613	340	171	149	474	136		225	.431	-.517	.465	-.293	.055	.757
Manikin	92	414	443	345	463	446	179	390	591	493	376	225		.694	.034	-.074	.068	-.304	.585
													LAMBDA	4.507	1.412	.993	.584	.390	
													PERCENT	37.56	11.77	8.27	4.87	3.25	

Item Name	Item No.	Varimax Rotation					Graphic Rotation (Orthogonal)					Biquartimin Rotation				
		Factors					Factors					Factors				
		1	2	3	4	5	A	B	C	D	E	1	2	3	4	5
Four buttons	55	.182	.912	-.132	-.069	-.043	-.13	-.13	.46	.52	.46	-.065	-.945	.017	-.034	.063
Three cube pyr.	69	.065	.844	-.177	-.049	-.406	.30	.14	.78	-.07	.10	.017	-.873	-.039	-.032	-.299
Two buttons	75	.032	.275	-.178	-.114	-.625	.27	-.01	.04	.11	.62	-.020	-.266	-.106	.098	-.565
Pict. puz. #2	76	.075	.281	-.849	-.166	-.005	.54	.03	.18	.52	.16	.165	-.140	-.916	.022	.202
Mare and foal	78	.101	.131	-.516	-.636	-.235	.51	.56	.16	.25	.34	.100	-.017	-.477	.595	-.067
Six cube pyr.	80	.212	.143	-.179	-.445	.157	.10	.46	.52	.22	-.09	-.077	-.092	-.118	.416	.241
Copy star	82	.361	.026	-.438	-.202	.033	.34	.19	.11	.41	.00	-.244	.065	.436	.125	.125
Action agent	85	.378	-.012	-.713	-.009	-.413	.77	-.02	.11	.38	.17	-.288	.137	-.751	-.100	-.295
Sequin form bd.	86	.600	.198	-.168	-.498	.006	.07	.50	.16	.58	.09	-.493	-.148	-.030	.453	.078
Decroly matching	88	.139	.135	-.552	-.135	-.323	.56	.05	.15	.28	.22	-.025	-.043	-.565	.057	-.199
Little pink tow.	89	.102	-.057	-.014	-.842	-.185	.08	.76	.03	.04	.31	.010	.107	.039	.904	-.114
Manikin	92	.527	.260	-.370	-.133	-.292	.29	.09	.01	.62	.29	-.460	-.204	-.286	.047	-.211
	LAMBDA	1.040	1.865	2.268	1.683	1.032										
	PERCENT	8.66	15.54	18.90	14.02	8.60										

TABLE 16. FACTOR ANALYSIS, STANFORD-BINET, FORM L
AGE LEVEL, 3 YEARS

Item Name	Item No.	Test Item Intercorrelations															Principal Axes (Unrotated)						
		2	III, 5	III, 6	III, 6-1	III, 6-2	III, 6-3	III, 6-4	III, 6-5	III, 6-6	IV, 2	IV, 4	IV, 5	IV, 6	IV-6, 3	IV-6, 5	1	2	3	4	5	6	h^2
Picture mem.	III,2		466	530	382	558	603	274	231	404	189	234	248	291	148	158	-.582	.262	-.250	.436	.147	.167	.709
Copy circle	III,5	466		466	536	205	513	695	593	573	-046	131	318	351	329	-151	-.650	.704	.331	.061	.104	-.173	1.000
Repeat 4 num.	III,6	530	466		292	657	214	479	309	534	396	610	350	258	469	392	-.749	-.062	-.198	-.015	.533	-.124	.904
Obey simple com.	III,6-1	382	536	292		378	230	284	207	064	230	354	167	086	066	-303	-.382	.473	-.300	-.179	.079	.106	.508
Pic. vocabulary	III,6-2	558	205	657	378		489	220	000	293	465	450	449	405	425	160	-.649	-.065	-.555	.166	.069	.061	.769
Comp. of sticks	III,6-3	603	513	214	230	489		488	269	238	435	311	426	447	315	111	-.629	.221	-.178	.299	-.480	.014	.796
Resp. to pict.	III,6-4	274	695	479	284	220	488		363	396	279	435	545	522	362	117	-.671	.219	.188	-.071	-.171	-.308	.663
Ident. obj. by use	III,6-5	231	593	309	207	000	269	363		422	094	459	244	455	210	-083	-.494	.293	.386	-.233	.002	.280	.611
Comprehension	III,6-6	404	573	534	064	293	238	396	422		204	487	425	668	567	454	-.707	-.127	.381	.092	.203	.048	.713
Naming obj, mem.	IV,2	189	-046	396	230	465	435	279	094	204		459	394	295	470	122	-.488	-.204	-.365	-.156	-.235	-.004	.493
Pict. ident.	IV,4	234	131	610	354	450	311	435	459	487	459		530	489	769	329	-.765	-.293	-.088	-.496	.020	.180	.958
Disc. of forms	IV,5	248	318	350	167	449	426	545	244	425	394	530		291	515	374	-.626	-.136	-.044	-.081	-.172	-.287	.530
Comp. II	IV,6	291	351	258	086	405	447	522	455	668	295	489	291		479	586	-.697	-.226	.353	.192	-.211	.278	.820
Pict. like & diff.	IV-6,3	148	329	469	066	425	315	362	210	567	470	769	515	479		330	-.683	-.303	.027	-.265	-.057	-.087	.640
Three commissions	IV-6,5	158	-151	392	-303	160	111	117	-083	454	122	329	374	586	330		-.380	-.713	.222	.363	.104	-.088	.852
														LAMBDA			5.791	1.792	1.287	.926	.772	.471	
														PERCENT			38.60	11.94	8.58	6.17	5.15	3.14	

Item Name	Item No.	Varimax Rotation						Graphic Rotation (Orthogonal)					Biquartimin Rotation					
		Factors						Factors					Factors					
		1	2	3	4	5	6	A	B	C	D	E	1	2	3	4	5	6
Picture mem.	III,2	-.061	.194	-.732	-.003	-.328	.157	.11	.31	.47	-.02	.59	.021	.028	-.921	-.229	.069	.048
Copy circle	III,5	.147	.877	-.308	.056	-.130	.408	-.14	.97	.26	-.01	.20	-.198	-.880	-.163	-.403	-.036	.242
Repeat 3 num.	III,6	-.202	.323	-.696	-.449	.248	.105	.13	.25	.73	.47	.18	.115	-.167	-.611	.199	-.531	-.127
Obey simple com.	III,6-1	.464	.220	-.413	-.190	-.084	.175	-.46	.30	.34	.14	.27	-.529	-.082	-.377	.161	-.048	.090
Pic. vocabulary	III,6-2	-.047	.008	-.692	-.460	-.272	-.057	.00	-.01	.46	.37	.64	-.031	.129	-.779	.346	.005	-.206
Comp. of sticks	III,6-3	-.004	.296	-.289	-.240	-.747	.098	.00	.37	-.11	.14	.79	-.044	-.352	-.339	.085	.582	.021
Resp. to pict.	III,6-4	-.057	.667	-.077	-.352	-.230	.179	-.03	.60	.00	.35	.28	-.008	-.840	.155	.112	.103	.014
Ident. obj. by use	III,6-5	.088	.328	-.046	-.139	-.039	.687	-.08	.68	.01	.26	-.07	-.120	-.144	.111	.096	.019	.713
Comprehension	III,6-6	-.470	.372	-.262	-.247	.018	.473	.45	.55	.21	.40	.07	.434	-.244	-.153	.010	-.164	.386
Naming obj., mem.	IV,2	.004	-.032	-.190	-.617	-.271	-.029	-.08	-.07	.02	.55	.41	-.073	.004	-.104	.655	.161	-.105
Pict. ident.	IV,4	-.072	.039	-.196	-.855	.051	.423	-.04	.20	.16	.91	.10	-.024	.108	.061	.919	-.156	.361
Disc. of forms	IV,5	-.193	.370	-.110	-.550	-.203	-.005	.08	.25	.00	.53	.33	.123	-.536	.083	.398	.069	-.163
Comp. II	IV,6	-.498	.146	-.115	-.280	-.385	.557	.49	.47	-.12	.41	.29	.478	-.011	-.086	.158	.296	.582
Pict. like & diff.	IV-6,3	-.248	.213	-.095	-.691	-.036	.213	.10	.22	.08	.73	.15	.173	-.243	.143	.628	-.072	.105
Three commissions	IV-6,5	-.885	-.036	-.111	-.234	-.023	.009	.82	-.08	.06	.34	.08	.889	.027	-.136	.050	-.082	-.046
LAMBDA		1.650	1.946	2.045	2.740	1.147	1.510											
PERCENT		11.00	12.97	13.63	18.27	7.65	10.06											

TABLE 17. FACTOR ANALYSIS, STANFORD-BINET FORM L
AGE LEVEL, 48 MONTHS

Item Name	Item No.	Test Item Intercorrelations												Principal Axes (Unrotated)				h^2
		1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	
Compar. of sticks	III-6,3		209	409	504	617	453	453	662	493	741	343	451	-.715	.017	.390	.069	.669
Naming objs., memory	IV,2	209		594	652	356	455	455	397	409	478	282	173	-.617	-.613	-.421	-.303	1.000
Pict. compl. (man)	IV,3	409	594		457	535	569	386	317	273	473	500	515	-.651	-.157	-.316	.223	.598
Comprehension II	IV,6	504	652	457		488	435	435	530	219	609	152	281	-.630	-.482	.107	.025	.641
Aesthetic compar.	IV-6,1	617	356	535	488		707	409	610	529	727	812	627	-.860	.220	-.062	.231	.844
Repeating 4 digits	IV-6,2	453	455	569	435	707		500	512	449	674	486	544	-.750	-.028	-.083	.179	.602
Pict. like & diff.	IV-6,3	453	455	386	435	409	500		349	198	527	233	305	-.543	-.265	.033	.168	.394
Materials	IV-6,4	662	397	317	530	610	512	349		561	809	364	621	-.780	.014	.385	-.205	.799
Three commissions	IV-6,5	493	409	273	219	529	449	198	561		570	672	607	-.687	.358	-.076	-.472	.829
Opposite analogies	IV-6,6	741	478	473	609	727	674	527	809	570		474	518	-.882	-.080	.279	-.013	.862
Paper folding	V,2	343	282	500	152	812	486	233	364	672	474		593	-.686	.493	-.398	.042	.874
Counting 4 objs.	V,6	451	173	515	281	627	544	305	621	607	518	593		-.693	.323	-.007	.028	.585
LAMBDA														6.116	1.234	.843	.528	
PERCENT														50.96	10.28	7.02	4.40	

Item Name	Item No.	Varimax Rotation				Graphic Rotation (Orthogonal)				Biquartimin Rotation			
		Factors				Factors				Factors			
		1	2	3	4	A	B	C	D	1	2	3	4
Compar. of sticks	III-6,3	-.270	-.109	.751	-.139	.74	.21	.15	-.18	-.092	.153	.818	.026
Naming objs., memory	IV,2	-.108	-.973	.103	-.239	.36	.57	-.04	.76	-.018	-1.032	.004	-.163
Pict. compl. (man)	IV,3	-.548	-.503	.208	.030	.60	-.02	.09	.50	-.573	-.300	.033	.242
Comprehension II	IV,6	-.089	-.589	.534	.028	.62	.38	-.19	.24	.091	-.410	.557	.140
Aesthetic compar.	IV-6,1	-.736	-.189	.491	-.160	.80	-.09	.41	.18	-.732	.105	.355	.120
Repeating 4 digits	IV-6,2	-.539	-.351	.429	-.073	.70	.04	.20	.27	-.502	-.107	.321	.150
Pict. like & diff.	IV-6,3	-.246	-.394	.413	.085	.56	.13	-.90	.21	-.155	-.184	.381	.225
Materials	IV-6,4	-.179	-.186	.760	-.393	.69	.45	.27	-.17	.042	-.034	.859	-.262
Three commissions	IV-6,5	-.379	-.130	.283	-.767	.36	.38	.75	.10	-.329	-.199	.243	-.644
Opposite analogies	IV-6,6	-.331	-.318	.773	-.231	.84	.34	.19	-.01	-.140	-.072	.809	-.040
Paper folding	V,2	-.836	-.084	.074	-.404	.46	-.18	.74	.34	-.971	.027	-.157	-.154
Counting 4 objects	V,6	-.569	-.047	.395	-.320	.58	.00	.50	.05	-.558	.125	.311	-.113
LAMBDA		2.593	2.035	2.916	1.177								
PERCENT		21.61	16.95	24.30	9.80								

TABLE 18. FACTOR ANALYSIS, STANFORD BINET, FORM L
AGE LEVEL, 60 MONTHS

Item Name	Item No.	Test Item Intercorrelations									Principal Axes (Unrotated)				h^2
		IV-6,5	V-1	V-4	V-5	VI-1	VI-2	VI-4	VI-5	VI-6	1	2	3	4	
Three commissions	IV-6,5		277	506	804	686	589	445	562	317	.732	.578	.258	.058	.940
Picture compl.	V-1	277		553	332	465	371	407	623	462	.576	-.387	.334	-.139	.613
Copy square	V-4	506	553		637	738	583	651	751	707	.834	-.203	.061	-.054	.744
Memory for sent.	V-5	804	332	637		776	525	750	717	586	.883	.360	-.223	-.236	1.000
Vocabulary	VI-1	686	465	738	776		583	651	683	550	.838	.109	.072	-.081	.725
Copy bead chain	VI-2	589	371	583	525	583		634	639	531	.727	.009	.040	.386	.679
Number concepts	VI-4	445	407	651	750	651	634		786	615	.822	-.106	-.306	.101	.790
Pict. like & diff.	VI-5	562	623	751	717	683	639	786		645	.883	-.181	.022	.013	.813
Maze tracing	VI-6	317	462	707	586	550	531	615	645		.718	-.277	-.131	-.021	.610
										LAMBDA	5.543	.787	.350	.247	
										PERCENT	61.58	8.75	3.89	2.75	

Item Name	Item No.	Varimax Rotation				Graphic Rotation (Orthogonal)				Biquartimin Rotation			
		Factors				Factors				Factors			
		1	2	3	4	A	B	C	D	1	2	3	4
Three commissions	IV-6,5	.157	.906	-.116	.285	.77	.40	.24	.34	-.004	.903	-.015	.119
Picture compl.	V-1	.747	.116	-.161	.124	-.08	.57	.54	-.04	-.830	.006	.074	-.029
Copy square	V-4	.591	.314	-.480	.256	.11	.80	.28	.11	-.547	.060	-.287	.137
Memory for sent.	V-5	.175	.737	-.663	.039	.60	.78	-.11	.07	-.013	.409	-.832	-.168
Vocabulary	VI-1	.422	.567	-.425	.211	.40	.71	.22	.15	-.332	.361	-.330	.048
Copy bead chain	VI-2	.276	.346	-.328	.612	.13	.58	.17	.56	-.122	.220	.068	.679
Number concepts	VI-4	.278	.262	-.718	.359	.07	.83	-.09	.27	-.126	-.100	-.579	.363
Pict. like & diff.	VI-5	.562	.336	-.524	.330	.11	.84	.25	.21	-.490	.067	-.298	.239
Maze tracing	VI-6	.460	.138	-.569	.236	-.05	.78	.10	.10	-.403	-.169	-.422	.186
	LAMBDA	1.822	2.118	2.110	.877								
	PERCENT	20.24	23.54	23.44	9.75								

Table 19. FACTOR INTERCORRELATIONS - BIQUARTIMIN ROTATIONS

	Factors	1.	2.	3.	4.
Cattell (3 months)	1.	--	.2915	-.2984	.3104
	2.	.2915	--	.3641	-.2504
	3.	-.2984	.3641	--	.0337
	4.	.3104	-.2504	.0337	--
Cattell (6 months)	1.	--	-.2013	-.4311	.2737
	2.	-.2013	--	-.3686	.4891
	3.	-.4311	-.3686	--	-.1660
	4.	.2737	.4891	-.1660	--

	Factors	1.	2.	3.	4.	5.	6.
California (6 months)	1.	--	.0376	-.0668	-.1063	-.1943	.0056
	2.	.0376	--	.0438	.0844	.0245	.0423
	3.	-.0668	.0438	--	.0868	-.1626	.1769
	4.	-.1063	.0844	-.0868	--	-.3795	.1877
	5.	-.1943	.0245	-.1626	-.3795	--	.0991
	6.	.0056	-.0423	.1769	.1877	.0991	--
California (12 months)	1.	--	-.1690	.0446	-.2044	.1809	.1175
	2.	-.1690	--	.1399	-.3027	-.0867	.0912
	3.	-.0446	-.1399	--	-.0989	.0085	.2914
	4.	-.2044	-.3027	-.0989	--	.1319	.0707
	5.	.1809	-.0878	.0085	.1319	--	.0841
	6.	.1175	.0912	.2914	.0707	.0841	--

Table 19 (Continued) FACTOR INTERCORRELATIONS - BIQUARTIMIN ROTATIONS

	Factors	1.	2.	3.	4.	5.	6.	7.	8.
Gesell (12 months)	1.	--	-.0607	-.0462	.1540	.1731	-.2263	-.1880	-.2868
	2.	-.0607	--	.0534	.0531	.1553	.1373	-.0813	-.1157
	3.	.0462	.0534	--	-.0217	.0104	-.0713	.4300	.1672
	4.	.1540	.0531	-.0217	--	-.1138	.0539	-.0310	.2097
	5.	.1731	.1553	.0104	-.1138	--	-.0099	-.1332	.0357
	6.	-.2263	.1373	-.0713	.0539	-.0099	--	-.1423	-.1172
	7.	-.1880	-.0813	.4300	-.0310	-.1332	-.1423	--	-.0873
	8.	-.2868	-.1157	.1672	.2097	.0357	-.1172	-.0873	--

	Factors	1.	2.	3.	4.	5.	6.
Merrill-Palmer (24 months)	1.	--	.1127	-.2630	.1238	.1009	.0985
	2.	.1127	--	.2482	.0227	.2757	-.1481
	3.	-.2630	.2482	--	.1409	.0467	.0488
	4.	.1238	.0227	.1409	--	.1185	-.0777
	5.	.1009	.2757	.0467	.1185	--	-.0199
	6.	.0985	-.1481	.0488	-.0777	-.0199	--
Merrill-Palmer (30 months)	1.	--	.2339	.0367	-.1783	.0497	
	2.	.2339	--	-.0726	-.0045	-.1450	
	3.	.0367	-.0726	--	.0921	.0140	
	4.	-.1783	-.0045	.0921	--	.3599	
	5.	.0497	-.1450	.0140	.3599	--	
Merrill-Palmer (36 months)	1.	--	.0023	-.0521	.3163	.2142	
	2.	.0023	--	-.2802	-.0193	.3185	
	3.	-.0521	-.2802	--	.2325	.1464	
	4.	.3163	-.0193	.2325	--	-.0687	
	5.	.2142	.3185	.1464	-.0687	--	

Table 19 (Continued) FACTOR INTERCORRELATIONS - BIQUARTIMIN ROTATIONS

	Factors	1.	2.	3.	4.	5.	6.
Merrill-Palmer (48 months)	1.	--	.1479	.0927	-.2006	-.0568	.1529
	2.	.1479	--	-.0035	-.0277	-.0765	-.2542
	3.	.0927	-.0035	--	-.0256	-.0431	-.2929
	4.	-.2006	-.0277	-.0256	--	-.0929	.1699
	5.	-.0568	-.0765	-.0431	-.0929	--	.1400
	6.	.0529	-.2542	-.2929	.1699	.1400	--
Merrill-Palmer (54 months)	1.	--	-.0571	-.2840	.1623	.2047	
	2.	-.0571	--	-.2784	.1043	-.0348	
	3.	-.2840	-.2784	--	.2172	-.2286	
	4.	.1623	.1043	.2172	--	.0186	
	5.	.2047	-.0348	-.2286	.0186	--	
	Factors	1.	2.	3.	4.	5.	6.
Stanford-Binet L (36 months)	1.	--	.0052	-.0347	-.2529	-.0330	.0008
	2.	.0052	--	-.4036	.2160	-.0216	.4351
	3.	-.0347	-.4036	--	.3206	.1292	.0921
	4.	-.2529	.2160	.3206	--	.0247	.0742
	5.	-.0330	-.0216	.1292	.0247	--	.1639
	6.	.0008	.4351	.0921	.0742	.1639	--
Stanford-Binet L (48 months)	1.	--	-.2230	.4602	-.2860		
	2.	-.2230	--	.2697	.3298		
	3.	.4602	.2697	--	-.0032		
	4.	-.2860	.3298	-.0032	--		
Stanford-Binet L (60 months)	1.	--	.0866	0.2099	.2920		
	2.	.0866	--	.3307	-.0952		
	3.	-.2099	.3307	--	.4623		
	4.	.2920	-.0952	.4623	--		

Table 20. SUMMARY OF FACTOR INTERPRETATIONS, ALLOWING COMPARISONS OF TESTS AND OF AGE LEVELS AS TO INTERPRETED FACTOR CONTENT

Intelligence Factor	Age-level (Months)															Total	
	Cat.		Calif.		Ges.		M-P					S-B(L)				
	3	6	6	12	6	12	24	30	36	48	54	36	48	60		
Cognitive																
CFU	X	X	X						X			X		X	6	
CFR				X							X				2	
CFS							X	X	X	X					4	
CFT								X		X	X			X	4	
CFI	X														1	
CMU							X								1	
CMR										X			X		2	
CMS											X				1	
CBU			X	X		X									3	24
Memory																
MFU														X	1	
MFR				X						X					2	
MFS									X						1	
MSS														X	1	
MMU													X		1	
MMR					X							X	X		3	
MBR				X		X									2	11
Divergent Prod.																
DFU	X					X									2	
DFR			X		X										2	
DFT		X													1	
DBR						X									1	6
Convergent Prod.																
NFU		X		X											2	
NFR	X				X				X	X					4	
NFS								X	X		X				3	
NMU						X									1	
NMI				X											1	
NBR						X		X							2	13
Evaluation																
EFU							X								1	
EFS					X							X			2	
EMR			X												1	
EMS		X	X		X										3	
EMT													X		1	8
Others																
Hand Dexty							X	X	X		X				4	
Gross Psychom						X	X								2	
Whole Body					X										1	
Locomotor					X										1	
Reflex					X										1	9
Total																71

Survey of Infant and Preschool Mental Testing

Name____________________ Organization________________

Official Title or Position______________________________

1. Do you include infants and (or) preschool children (under six years of age) in your clientele?

   Yes________ No________ Approximate number per year________

   (If your answer to this question is "no," disregard the rest of this questionnaire, and please return it in the stamped, addressed envelope provided.)

2. Services offered by your organization for infants and preschool children: (Please check)

a. Testing

___ Intelligence, developmental
___ Diagnostic
___ Pre-adoption
___ Pre-admission
___ Physical handicaps
___ Other ____________
___ ____________

b. Treatment

___ Psychotherapy
___ Remedial
___ Other ____________
___ ____________

c. Educational

___ Nursery School
___ Residential School
___ Day care
___ Other ____________
___ ____________

d. Research

___ Anoxia
___ Autism
___ Other ____________
___ ____________

e. Medical

___ Pediatric
___ Psychiatric
___ Out-Patient
___ Other ____________
___ ____________

3. Types of cases (infants and children under six) handled (check):

Approx. no. per year

_____ Normal.................................. _____
_____ Retarded................................ _____
_____ Behavior difficulties................... _____
_____ Verbal disability....................... _____
_____ Physically handicapped.................. _____
_____ Physically disabled..................... _____
_____ Brain damaged........................... _____
_____ Autistic................................ _____
_____ Psychotic............................... _____
_____ Other (specify)......................... _____

4. Please check below the tests that are in use in your organization with children under six:

Estimated Average Number Per Year

_____ Ammons Picture Vocabulary............... _____
_____ Bender-Gestalt.......................... _____
_____ Blackie Test............................ _____
_____ Buhler Infant Test...................... _____
_____ California Infant Scale................. _____
_____ California Pre-School Scale............. _____
_____ C A T................................... _____
_____ Cattell Infant Intelligence Scale....... _____
_____ Columbia................................ _____
_____ Goodenough Draw-a-Man................... _____
_____ Frazer Visual Perception Test........... _____
_____ Gesell Developmental.................... _____
_____ Grace-Arthur............................ _____
_____ Griffith's Abilities of Babies.......... _____
_____ Kuhlman-Binet........................... _____
_____ Leiter International Performance Scale.. _____
_____ Merrill-Palmer Scale of Mental Tests.... _____
_____ Minnesota Pre-School Scale.............. _____
_____ Picture Story Test...................... _____
_____ Picture Vocabulary...................... _____
_____ Randall's Island........................ _____
_____ Raven Matrices.......................... _____
_____ Rorschach............................... _____
_____ Shotwell, Dingman & Tarjan Number
Concept Scale........................... _____
_____ Smith non-verbal........................ _____
_____ Stanford-Binet.......................... _____
_____ Valentine Intelligence Test............. _____
_____ Verbal Discrimination Test.............. _____
_____ Vineland Social Maturity................ _____
_____ W I S C................................. _____
_____ Others (specify) _____
_____ _____
_____ _____

5. Which of the above tests do you consider to be of special value and in what way?

6. What are the inadequacies or limitations of the tests you are using?

7. In what ways do you feel that the present available mental tests could be improved?

8. If a new and more modern mental test for infants and preschool children were to be developed which would meet more adequately the needs in this field, would your organization be willing to cooperate in the acquisition of the necessary data provided the examiners used are trained and reimbursed by the research project?

Yes ______ No ______

Questionnaire submitted by Drs. Leland H. Stott and Rachel S. Ball. This research is being carried out as a project sponsored by the Cooperative Research Branch of the Office of Education, U. S. Department of Health, Education and Welfare.

Please return questionnaire in enclosed stamped envelope to:

Dr. Leland H. Stott
The Merrill-Palmer Institute
71 East Ferry Avenue
Detroit 2, Michigan